DUPLICATE ALSO

Forecasting Research

Forecasting Research Division
Scientific Paper No. 49

Some Applications of Transformation Theory in Mechanics

by

M.J. Sewell

19th October 1997

Meteorological Office
London Road
Bracknell
Berkshire
RG12 2SZ
United Kingdom

ORGS UKMO F

National Meteorological Library
FitzRoy Road, Exeter, Devon. EX1 3PB

Some Applications of Transformation Theory in Mechanics

by **M.J. Sewell**

Department of Mathematics, University of Reading

and

Numerical Weather Prediction Development Division, Meteorological Office, Bracknell

19th October 1997

This chapter has been prepared for the Proceedings of the Isaac Newton Institute Programme on The Mathematics of Atmosphere and Ocean Dynamics, held in Cambridge from July to December 1996. The Proceedings will be published by Cambridge University Press.

Contents

1.	Introduction	3
2.	Duality	5
3.	Catastrophes	27
4.	Upper and lower bounds	35
5.	Lift and contact transformations	39
6.	Quartets of Legendre transformations	45
7.	Flux of a jacobian	53
8.	Canonical transformations	55
9.	Central orbits and semi-geostrophic theory	68
10.	Hamiltonian structures	77
	References	85

Abstract

A review is given of some of the properties of duality, including Legendre duality and other types, and of lift transformations and canonical transformations. Applications are mentioned in several branches of mechanics. A straightforward style is adopted, so that the paper is accessible to a wide readership.

Developments in the semi-geostrophic theory of meteorology in the last fifteen years have prompted this review, but it draws upon earlier work in, for example, plasticity theory, gas dynamics, catastrophe theory, hamiltonian mass-point mechanics, and the theory of maximum and minimum principles.

Singularities need to be described in transformation theory, and the swallowtail catastrophe is one such example. The convexified swallowtail is used to picture an atmospheric front. The intimate relation between lift transformations and hamiltonian structures is described. New exact solutions in a semi-geostrophic central orbit theory are given.

1. Introduction

The purpose of this article is to give a widely accessible review of some properties of certain transformations of functions which are used, in mechanics in particular, to convert equations into alternative versions which are more convenient, recognisable or amenable to solution than the original ones. We focus upon duality, both of Legendre and of other types, upon lift transformations (which include contact transformations as a special case) and canonical transformations, and certain other coordinate transformations. Some topics are treated in the lowest practicable dimensional case, because this allows ready progress through a sequence of ideas whose diversity and interconnections can be appreciated easily. Thus a fabric is woven whose pattern can be seen quickly. Generalization to higher dimensions will be straightforward in some cases, and less so in other, where some desirable new research directions may be prompted, but we have not space to pursue such here.

Transformation theory is a subject which has interested me for forty years, and it has a time-honoured and large literature. Contexts in which I have used it vary from the Legendre transformation in plasticity theory and gas dynamics in earlier papers, through a study of singularities, represented by the swallowtail in catastrophe theory, to, most recently, canonical and lift transformations in meteorology. This review was triggered by the latter interest. There has not been space here to discuss meteorology in full detail. That subject is well catered for by most of the other review papers in this Volume. However, some features of the so-called f-plane semi-geostrophic theory are illustrated in several places. Also, as a modest prototype of that theory, I include a new comparison in straightforward language of the Newtonian and semi-geostrophic motions of a particle under a central force, as seen from a rotating reference frame.

Notwithstanding the age and size of transformation theory as a subject, I have felt that a fresh view might be helpful. Some new results are included here, and some relatively recent ones are represented, such as the picturing of an atmospheric front via a convexified swallowtail.

The plan of the paper is as follows. §2 contains a discussion of duality from a symmetric starting point, which allows the contrast to be made between the very common Legendre duality, and less familiar other types. Applications mentioned include some in vortex dynamics and some in plasticity theory, more classical types such as the circle of Appolonius, and phase and slowness surfaces in wave mechanics, and recent results about duality in meshes such as may be appropriate to numerical analysis. In §3 some brief properties of the elementary catastrophes are recounted, because of their central role in classifying singularities in transformations. A classroom experiment demonstrating instability is described. A brief mention is made in §4 of the structure of the theory of upper and lower bounds. This is in the context of a wide class of problems governed by systems of equations or inequalities generated by a saddle functional and an associated hamiltonian structure, with an example from gas dynamics. A straightforward account of some salient properties of lift transformations is given in §5. A particular example inspired by the so-called f-plane (a local tangent plane to the Earth) semi-geostrophic theory of meteorology is included. Quartets of Legendre transformations are exhibited in §6. This topic can be viewed as originating in classical thermodynamics, but here we give examples associated with the modelling of atmospheric fronts in semi-geostrophic theory. Jumps in air temperature and wind velocity at a front are represented by gradient jumps in geopotential surfaces. In §7 a general formula for calculating a parametric derivative of a jacobian is recorded. This applies in particular to potential vorticity. A comparison of five different definitions of a canonical transformation is described in §8. Several viewpoints are offered, and their relation to hamiltonian mechanics is

included. In §9 we study the classical central orbit problem from the standpoint of a rotating plane, and compare it with its semi-geostrophic replacement. We find that the solutions of the latter are circles, which are the envelope of the classical Newtonian rotating conic solutions under the inverse square law. This result for ordinary differential equations illustrates how drastic is the filtering imposed by the semi-geostrophic approximation, although it is not a direct comment on the properties of the partial differential equations of meteorology *per se*. In §10 we exhibit the hamiltonian structure of the equations solved in §9, as a prototype of the corresponding structures expected in meteorology and shallow water theory. The semi-geostrophic approximation replaces one hamiltonian system with another hamiltonian system of half the size. A general theorem about canonical transformations is also given in §10.

Lucidity is something which has to be worked at in applied mathematics. In striving for it here I have adopted the theorem and proof style of exposition. This is more customary in pure than in applied mathematics, but I have come to believe that it has some advantages in sharpening, even for oneself, what one wishes to say.

2. Duality

Duality is a widely used concept in mathematics, and the starting point for studying it is a matter of taste. Here we begin by supposing that a real scalar function $F(x_i, y, X_i, Y)$ of $2(n+1)$ real variables is given, with $i = 1, 2, ..., n$. The equation

$$F(x_i, y, X_i, Y) = 0 \tag{1}$$

is set down. This implicit relation allows the following definitions.

For each given point in x_i, y space, which we call a *pole*, (1) specifies a surface in X_i, Y space, which we call a *polar*. Alternatively, for each given point in X_i, Y space, also

called a pole, (1) specifies a surface in x_i, y space, also called a polar. In each case, we can say that the pole and polar are *dual* to each other. These coordinates need not be cartesian.

If the pole is now allowed to move over a specified curve or surface of dimension $k \le n$, the polar will move and envelop another surface whose dimension K will depend on the given k and n. We can say that two such surfaces are also dual to each other. We might expect that $k + K \ge n$. Care is required in discussing envelopes because, as Fig. 1 illustrates, an envelope can consist of disjoint parts having different dimensions. Some such parts may have $K > n - k$ and others $K = n - k$, for the same value of $n - k$.

Fig. 1 is an example which does use cartesian coordinates with $n = 1$, and the pole moves on $y = \frac{1}{2}(1 + x^2)$ (dashed). The polars all pass through the origin, and also envelop

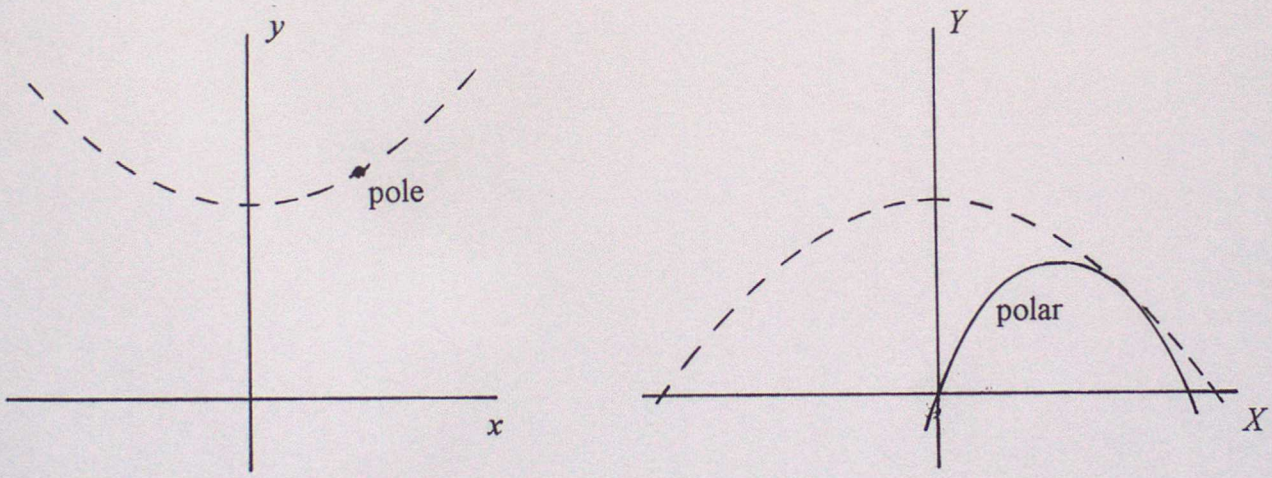


Fig. 1 Pole and polar, and dual curves (dashed) for $F = xX - yX^2 - Y$.

$Y = \frac{1}{2}(1 - X^2)$ (also dashed). These polars can be viewed as a family of trajectories, of a projectile from the origin under gravity, whose envelope is the parabola of safety from them.

The definition of duality is linked to the choice of the particular function $F(\,)$, and we can recognise that explicitly by speaking of F-duality. Another function $G(\,)$ would specify G-duality. The following two questions arise immediately.

1. Under what circumstances can two types of duality co-exist?

2. When they co-exist, under what circumstances are they equivalent?

We define co-existence here, and we discuss equivalence later.

Evidently we can say that the existence of G-duality implies that of F-duality if there exists a transformation

$$\begin{aligned} w_i &= w_i(x_j, y, X_j, Y) \quad , \quad p = p(x_j, y, X_j, Y) \\ W_i &= W_i(x_j, y, X_j, Y) \quad , \quad P = P(x_j, y, X_j, Y) \end{aligned} \tag{2}$$

such that

$$G(x_i, y, X_i, Y) = F(w_i, p, W_i, P) \ . \tag{3}$$

For then $G = 0$, over any G-domain where (2) exists, implies $F = 0$ over the corresponding F-domain. If the inverse of (2) exists, then the existence of F-duality implies that of G-duality in a similar sense.

Legendre duality is a familiar type, defined by $F = 0$ with the particular choice

$$F(x_i, y, X_i, Y) = x_i X_i - y - Y \ . \tag{4}$$

Many properties of Legendre duality are described by Sewell (1987, Chapter 2; the transcriptions $X_i \to y_i$ and $y \to X$ are needed to convert the present notation to that used there). Any bilinear function, such as (4), or $F(x_i, y, X_i, Y) = x_i X_i - yY$ which we introduce in (43), implies plane polars, and these can be especially helpful in computational applications. As we have just seen, for some specific purposes it is desirable to change the symbols used for the arguments of a given function, such as $F(\ )$ and $G(\ )$. In other words, it is flexible to regard such arguments as denoted by dummy variables.

Another type of duality, which offers itself in the semi-geostrophic theory of meteorology with cartesian coordinates, has

$$G(x_i, y, X_i, Y) = Y - y - \tfrac{1}{2}(x_i - X_i)^2 \tag{5}$$

(see Sewell and Roulstone, 1994, p. 509). However, it is readily seen that the invertible transformation

$$w_i = x_i\ ,\ p = \tfrac{1}{2}x_i^2 + y \quad ,\quad W_i = X_i\ ,\ P = \tfrac{1}{2}X_i^2 - Y\ , \tag{6}$$

used with the particular functions (4) and (5), implies the co-existence of Legendre duality and semi-geostrophic duality, because (3) holds. In the meteorology literature this fact appears in another guise with more hypotheses wherein, with cartesian coordinates, a geopotential function $y(x_i)$ proportional to physical height is employed to construct a 'modified geopotential' function $p(x_i) = \frac{1}{2}x_i^2 + y(x_i)$, with $n = 2$ (cf. Cullen and Purser, 1984, equations (2.6) and (3.11)).

It seems intuitive that Legendre duality and semi-geostrophic duality not only co-exist, but are equivalent, and we seek to articulate a criterion for equivalence. Define the transformation (2) to be *separated* if it has the particular form

$$w_i = w_i(x_j, y)\ ,\quad p = p(x_j, y)\ ,\quad X_i = X_i(W_j, P)\ ,\quad Y = Y(W_i, P)\ . \tag{7}$$

Then we can say unambiguously that any particular choice of pole in x_i, y space implies a pole in w_i, p space; and that any particular choice of polar in X_i, Y space implies a polar in W_i, P space. The following result applies.

Theorem 1

If (7) is invertible and assures the co-existence of F-duality and G-duality in the sense of (3), then for any initial choice of pole, the transform of its polar is the same as the polar of its transformed pole.

Proof

We explain and prove the result in terms of the closure of the diagram in Fig. 2. Choose any particular pole $x_i = a_i, y = b$ and calculate (vertical arrow) its G-polar. Use the

transformation (horizontal arrows) to calculate the transformed pole, and also the transform of the G-polar. This latter is the same, by the particular application

$$G\big(a_i, b, X_i(W_j, P), Y(W_j, P)\big) = F\big(w_i(a_j, b),\, p(a_j, b),\, W_i,\, P\big) \tag{8}$$

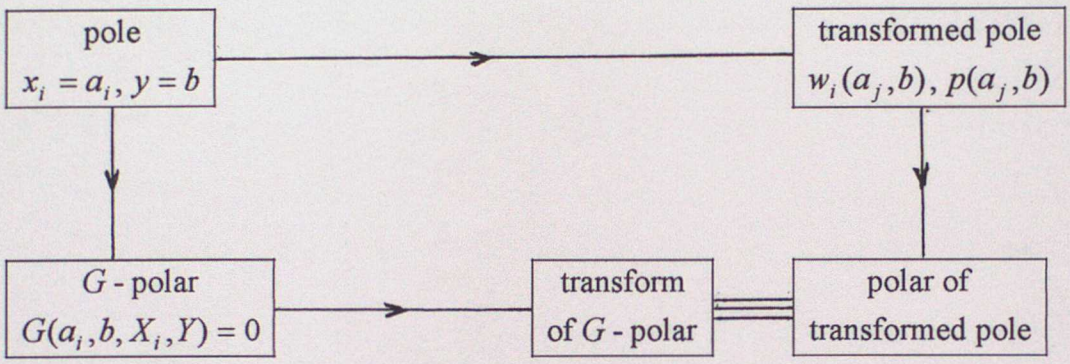


Fig. 2 Proof of Theorem 1.

of the co-existence equation (3), as the F-polar calculated (vertical arrow) for the transformed pole, putting both sides of (8) to zero.

□

The two functions of W_i and P in (8) are the same, before either is required to be zero. We have not actually used the invertibility of (7) in reaching (8), but we should do so if we had started with a pole in w_i, p space instead of in x_i, y space.

The Legendre/semi-geostrophic example may be followed through as an illustration of the calculation in Fig. 2. Using (6) as an illustration of (7), the pole $x_i = a_i,\ y = b$ has the G-polar

$$G(a_i, b, X_i, Y) = Y - b - \tfrac{1}{2}(a_i - X_i)^2 = 0 \tag{9}$$

whereas the transformed pole $w_i = a_i,\ p = \tfrac{1}{2}a_i^2 + b$ has the F-polar

$$F(a_i, \tfrac{1}{2}a_i^2 + b, W_i, P) = a_i W_i - P - \tfrac{1}{2}a_i^2 - b = 0\ . \tag{10}$$

Using $X_i = W_i, Y = \frac{1}{2}W_i^2 - P$ confirms that (9) and (10) contain the same functions, as an illustration of (8).

One can take the view that such a property as Theorem 1 expresses, which presupposes a separated transformation, would be necessary and sufficient to serve as the definition of equivalence for co-existing dualities. On that basis, dualities which co-exist via transformations (2) which are *not* separated in the sense of (7), but *mixed* instead, might not be equivalent. We shall give an example of such in (47).

Another viewpoint is the following. For $n = 1$ and constants a, b, c the equation

$$\tfrac{1}{2}(ac \pm 1)ax^2 + abx \mp y - Y + \tfrac{1}{2}c(X - ax)^2 + \left[(ac \pm 1)x + b\right](X - ax) = 0 \qquad (11)$$

includes both Legendre duality (with $a = b = c = 0$ and the upper choice of sign) and semi-geostrophic duality (with $a = c = 1$, $b = 0$ and the lower choice of sign). The genesis of (11) is via a study of lift transformations for general n (Sewell and Roulstone, 1994, Theorem 8 and equation (71)). We shall explain that viewpoint presently.

Vortex dynamics is a context which provides another set of examples in which some types of semi-geostrophic duality imply Legendre duality, but in this case using coordinates which are not cartesian. Purser (1993, §5) gave an axisymmetric vortex theory whose duality is defined, in the first instance, by setting to zero the function

$$G(r,z,y,R,Z,Y) = Y - y - \frac{f^2}{8}\left(\frac{R^4}{r^2} + r^2\right) + \frac{f^2}{4}R^2 + Zz \ . \qquad (12)$$

Here f is the constant Coriolis parameter, r and z are cylindrical polar coordinates of radius and height of axisymmetric motion in the physical space, and R and Z can be thought of as having similar roles in a dual space. There exists a transformation

$$p = \frac{f^2}{8}r^2 + y \quad , \quad P = -f^2\frac{R^2}{4} - Y \ , \qquad (13)$$

of type $(7)_2$ with $(7)_4$ for which $G = \Gamma(r,z,p,R,Z,P)$ where

$$\Gamma = -\frac{f^2 R^4}{8r^2} + zZ - p - P \ . \tag{14}$$

This is not yet bilinear as the Legendre form (4) requires. Another mapping

$$\hat{r} = -\frac{1}{2r^2} \quad , \quad \hat{R} = f^2 \frac{R^4}{4} \tag{15}$$

of type $(7)_1$ with $(7)_3$ is sufficient to reach $\Gamma = F(\hat{r}, z, p, \hat{R}, Z, P)$ where

$$F = \hat{r}\hat{R} + zZ - p - P \ . \tag{16}$$

Equations (13) and (15) are given in Purser (*op.cit.*, equations (5.7)). The first of (15) can be recognized in Shutts *et al.* (1988, equation (4)), where it is motivated by the need for an $\hat{r}$-gradient of a geopotential modified by the addition of a centrifugal term. Summarizing, the mapping (13) with (15) is an example of a separated transformation (7) which converts the $G = 0$ duality defined by (12) and (13) for axysymmetric vortex dynamics into Legendre duality.

Non-axisymmetric vortex theory on an f-plane requires the introduction of angles θ and Θ representing longitude in the physical and dual spaces. Purser (1996) gives a theory in which $F = 0$ from (16) is replaced by the requirement

$$\hat{r}\hat{R}\cosh\left(\sqrt{8}(\theta - \Theta)\right) + zZ - p - P = 0 \ . \tag{17}$$

This becomes bilinear and the starting point for a theory of Legendre duality, in terms of new variables defined by

$$\hat{x} = \hat{r}\cosh\sqrt{8}\,\theta \ , \quad \hat{y} = -\hat{r}\sinh\sqrt{8}\,\theta \ , \quad \hat{X} = \hat{R}\cosh\sqrt{8}\,\Theta \ , \quad \hat{Y} = \hat{R}\sinh\sqrt{8}\,\Theta \ , \tag{18}$$

because then

$$\hat{x}\hat{X} + \hat{y}\hat{Y} + zZ - p - P = 0 \ . \tag{19}$$

Purser also indicates how similar methods apply to semi-geostrophic theory of three-dimensional zonal flow for a hemispherical vortex.

The definition of duality from the starting point chosen here is a discrete one. Poles map into polars, whether the latter are plane or not. Discreteness is therefore fundamental. When the pole is allowed to move, that will be on some specified surface or lower dimensional object such as a curve, which may or may not be smooth or single-valued or even continuous. Such features will be reflected in peculiarities of the envelope of the polar. This is true in particular for Legendre duality, which has a plane polar.

If rather strong assumptions are made about the surface on which the pole moves, conclusions represented by the following Theorem emerge.

Theorem 2

(a) When the pole moves on a surface having the parametric description

$$x_i = x_i(u_\alpha) \quad , \quad y = y(u_\alpha) \tag{20}$$

in terms of differentiable functions (on the right) of $m \le n$ parameters u_α, (1) has the properties

$$\frac{\partial F}{\partial x_i}\frac{\partial x_i}{\partial u_\alpha} + \frac{\partial F}{\partial y}\frac{\partial y}{\partial u_\alpha} = 0 \ . \tag{21}$$

(b) When the polar envelops a surface

$$X_i = X_i(v_\beta) \quad , \quad Y = Y(v_\beta) \tag{22}$$

described by differentiable functions of m other parameters v_β related to the u_α by differentiable functions $u_\alpha(v_\beta)$ for $\alpha, \beta = 1, \ldots, m$, (1) also has the properties

$$\frac{\partial F}{\partial X_i}\frac{\partial X_i}{\partial v_\beta} + \frac{\partial F}{\partial Y}\frac{\partial Y}{\partial v_\beta} = 0 \ . \tag{23}$$

Proof

(a) We regard the envelope as being defined by eliminating the u_α between $F(x_i(u_\alpha), y(u_\alpha), X_i, Y) = 0$ and the m equations $\partial F / \partial u_\alpha = 0$. Then (21) is immediate by the chain rule.

(b) The envelope also satisfies

$$F\left[x_i\left(u_\alpha(v_\beta)\right), y\left(u_\alpha(v_\beta)\right), X_i(v_\beta), Y(v_\beta)\right] = 0,$$

so (23) follows from $\partial F / \partial v_\beta = 0$ for each β.

□

Theorem 3

(a) Under the hypothesis of Theorem 2(a), Legendre duality has the property

$$X_i \frac{\partial x_i}{\partial u_\alpha} = \frac{\partial y}{\partial u_\alpha} \quad . \tag{24}$$

Where (20) can be written as $y = y(x_i)$, (24) becomes

$$X_i = \frac{\partial y}{\partial x_i} \quad . \tag{25}$$

(b) Under the hypothesis of Theorem 2(b), Legendre duality has the property

$$x_i \frac{\partial X_i}{\partial v_\beta} = \frac{\partial Y}{\partial v_\beta} \quad . \tag{26}$$

Where (22) can be written as $Y = Y(X_i)$, (26) becomes

$$x_i = \frac{\partial Y}{\partial X_i} \quad . \tag{27}$$

Proof

The application to $F = x_i X_i - y - Y$ of (21) gives (24), and of (23) gives (26). The familiar formulae (25) and (27) are recovered when it is sufficient to use x_i and X_i as the parameters u_α and v_α respectively.

□

It is when (25) and (27) both apply that the Legendre transformation can be described as a dual gradient transformation. We shall see, however, that there are cases when one applies but not the other (cf. Theorem 5). This is why we have presented each of Theorems 2 and 3 in two separate parts.

A discrete example of Legendre duality for $n = 1$ is shown in Fig. 3. The pole in x, y space moves on a piecewise smooth curve consisting of straight segments, a parabola, and points joining them. The polar in X, Y space envelops another piecewise smooth curve consisting of corresponding points, another parabola, and straight line segments, lettered to indicate the correspondence. The curve in X, Y space can also be regarded as the locus of a pole there, whose polar in x, y space envelops the curve shown there. The gradients satisfy $X = dy/dx$ and $x = dY/dX$ by (25) and (27), and they are related by the step function shown in Fig. 3, with an inclined segment corresponding to the parabolae.

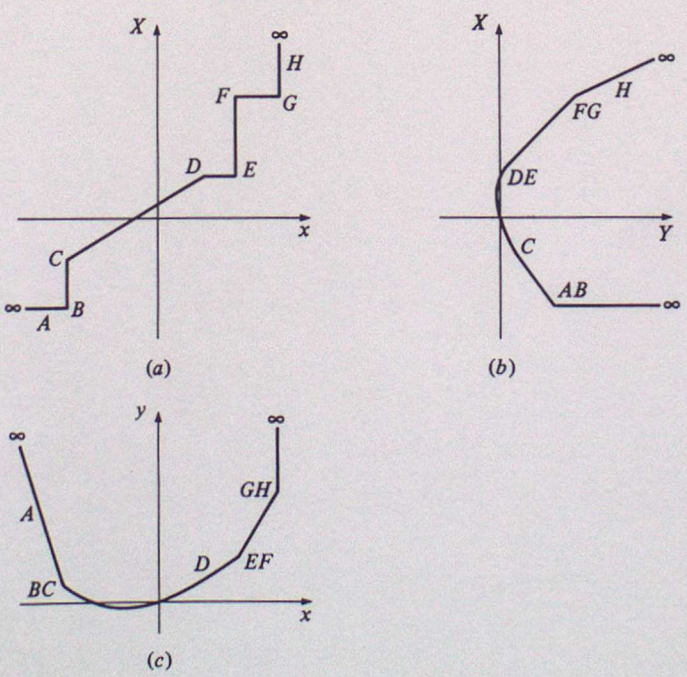


Fig. 3 Discrete example of Legendre duality.

The dual functions $y(x)$ and $Y(X)$ illustrated in Fig. 3 happen to be weakly convex, so that their gradients are related by a weakly monotonic function. It is important to note, however, that convexity is not at all necessary in the definition of duality. When convexity is present, it frequently indicates a stability property in an associated application.

Fig. 4 illustrates a particular way in which convexity can fail, in this case for smooth functions with $n = 1$. The curve $X = X(x)$ having a local maximum could represent a 'response curve' relating, for example, stress and strain in an elastic solid, pressure and radius

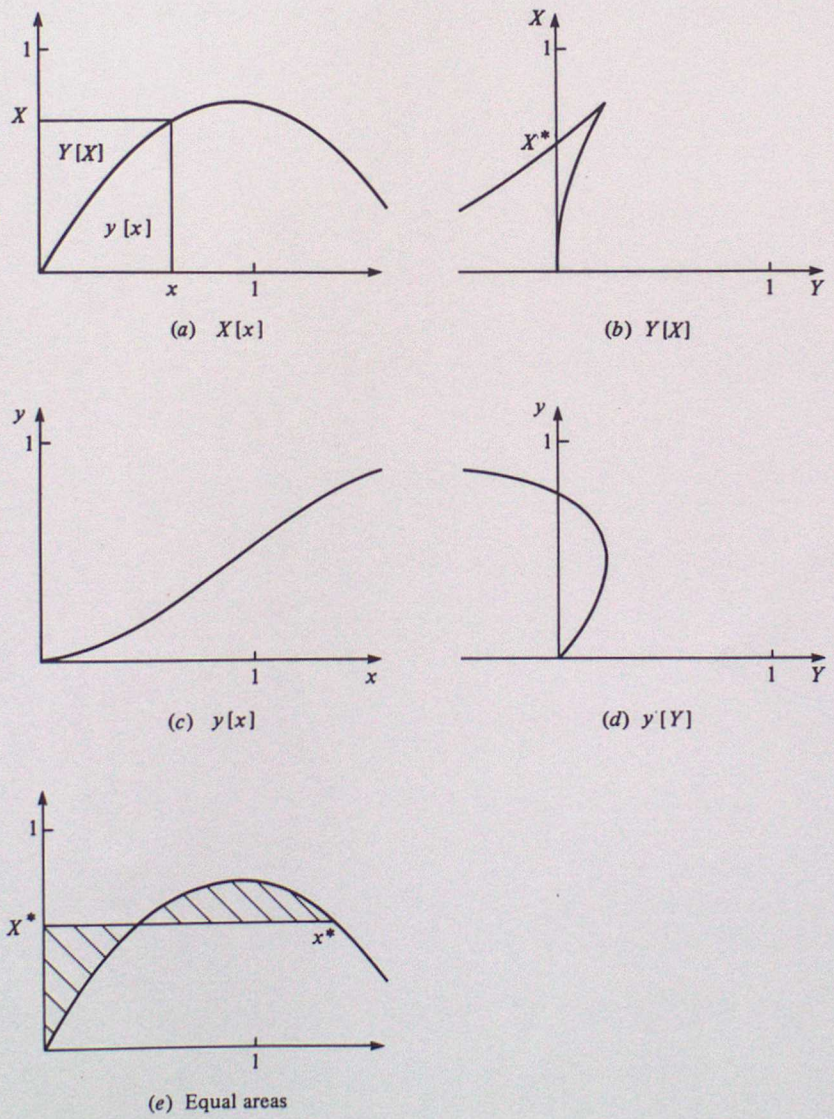


Fig. 4 Non-convex Legendre dual functions $y(x)$ and $Y(X)$.

in a balloon (Sewell, 1997, Chap. 8), or flow stress and speed in a fluid (Sewell, 1987, Fig.5.3). The areas $y(x)$ and $Y(X)$ between the curve and the two axes satisfy $xX = y + Y$, and they are Legendre dual functions. The local maximum of $X(x)$ induces an inflexion in $y(x)$ and a cusp in $Y(X)$. The particular value X^* (say) where $Y(X^*) = 0$ locates the height of a horizontal intercept of the response curve where equal areas (shaded) lie between the curve and the intercept.

There is a practical aspect of Legendre duality which arises when one of the functions $y(x_i)$ or $Y(X_i)$ is regarded as the starting point from which to calculate the other one, as follows.

Theorem 4

If $y(x_i)$ is given, the determinantal condition

$$\left| \frac{\partial^2 y}{\partial x_i \partial x_j} \right| \neq 0 \quad \text{or} \quad \pm\infty \tag{28}$$

is sufficient, but not necessary, to allow the local inversion of $X_i = \partial y / \partial x_i$, so that the inverse $x_i = x_i(X_j)$ can be calculated and inserted into

$$x_i X_i = y + Y \ . \tag{29}$$

This determines the Legendre dual function

$$Y(X_i) = X_i\, x_i(X_j) - y\left[x_i(X_j)\right] \tag{30}$$

having the property $x_i = \partial Y / \partial X_i$. Also,

$$\frac{\partial^2 Y}{\partial X_i \partial X_k} \frac{\partial^2 y}{\partial x_k \partial x_j} = \delta_{ij} \tag{31}$$

and therefore, by taking determinants,

$$\left| \frac{\partial^2 Y}{\partial X_i \partial X_k} \right| \left| \frac{\partial^2 y}{\partial x_k \partial x_j} \right| = 1 \ . \tag{32}$$

The proof is by the inverse function theorem.

The condition (28) often fails at isolated places, such as the inflexion on $y(x)$ and the cusp on $Y(X)$ illustrated in Fig. 4, but such *isolated singularities*, as they may be called, do not inhibit the global calculation of the Legendre dual function, even though they may terminate a range of convexity or single-valuedness. Two specific illustrations for $n = 1$ are shown in Fig. 5, with $a = -2$ and $a = +1$ in the functions

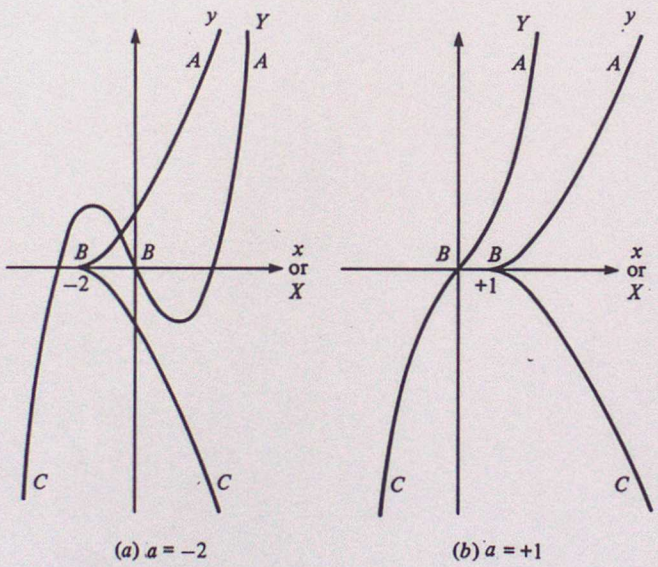


Fig. 5 Isolated singularities at inflexions and cusps

$$Y(X) = \tfrac{1}{3} X^3 + aX \quad \text{and} \quad y(x) = \pm \tfrac{2}{3} (x-a)^{3/2} \ . \tag{33}$$

The properties in (28)-(32), and the isolated singularity, can be explored explicitly with this example.

There are some circumstances, however, when the hypothesis (28) fails not just locally, but globally, and then we must have recourse to the more basic definition of duality given earlier, of which (25) and (27) in Theorem 3 are just one particular consequence. This can be illustrated with homogeneous functions.

Theorem 5

If $y(x_i)$ is a twice differentiable homogeneous function of degree r, then the definitions

$$X_i = \frac{\partial y}{\partial x_i} \quad \text{and} \quad Y = x_i X_i - y \tag{34}$$

imply that

(a) $Y \equiv 0$ if $r = 1$; (35)

(b) a Legendre dual function $Y(X_i)$ which is also homogeneous of order r cannot exist unless $r = 0$ or $r = 2$.

Proof

Euler's theorem for homogeneous functions gives

$$x_i \frac{\partial y}{\partial x_i} = r\, y \ . \tag{36}$$

Therefore

$$Y = x_i \frac{\partial y}{\partial x_i} - y = (r-1)y \ . \tag{37}$$

This proves (a).

When $r \neq 1$ and we assume that $Y(X_i)$ exists with the property $x_i = \partial Y / \partial X_i$,

$$X_i \frac{\partial Y}{\partial X_i} = Y + y = \frac{rY}{r-1} \ . \tag{38}$$

This is rY where $Y \neq 0$ if and only if $r(r-2) = 0$, which proves (b).

□

We note that the definition $(34)_1$ required to prove (a) is justified by (25). In other words, the hypothesis (20) of Theorem 2(a) applies here in the particular form $y = y(x_i)$ required by Theorem 3(a), so that (25) is established.

For $r = 1$, however, the additional hypotheses needed to ensure that (22) can be written as $Y = Y(X_i)$ satisfying (27) are evidently not available. Also, Theorem 4, which is the usual practical basis for the calculation of Legendre dual functions, is not available either because (28) fails globally, and not merely locally. This can be seen by differentiating (36) again, which leads to

$$x_i \frac{\partial^2 y}{\partial x_i \partial x_j} = (r-1)\frac{\partial y}{\partial x_j} \quad \text{and} \quad x_i \frac{\partial^2 y}{\partial x_i x_j} x_j = r(r-1)y \ . \tag{39}$$

When $r = 1$ (39) shows that at least one eigenvalue of the hessian is identically zero, namely that eigenvalue associated with a radial eigenvector.

Examples indicating what can happen are provided by planar and conical functions such as

$$y = c_i x_i \quad \text{and} \quad y = (c_i x_i^2)^{1/2} \tag{40}$$

for constant c_i. In such cases we have the options of looking to see if anything survives from Theorem 4 in a subspace orthogonal to the radial eigenvector, or of going back to the more basic principles for the calculation of envelopes of polars indicated at the beginning of this Section. If the pole roams over the plane $y = c_i x_i$, then the polar $Y + y = X_i x_i$ becomes $Y = (X_i - c_i)x_i$ which, for all x_i, envelops the single point

$$X_i = c_i \quad , \quad Y = 0 \tag{41}$$

only, as (34) also show. This not only illustrates (35), but represents the dual function $Y(X_i)$ as an *accumulated singularity* (Sewell, 1987), where $Y(c_i) = 0$ and $Y(X_i \neq c_i)$ is undefined.

If the pole roams over the cone $y = (c_i x_i^2)^{1/2}$, then $yX_i = c_i x_i$ (no sum) and

$Y = \left(\frac{X_i^2}{c_i} - 1 \right) y$ (summing) from (34), so for all x_i and therefore for all y the polar envelops the 'conic section'

$$\frac{X_i^2}{c_i} = 1 \quad , \quad Y = 0 \ . \tag{42}$$

This again illustrates (35), and additionally shows where the dual function is accumulated in the base plane. For example, the Legendre dual of an elliptic or hyperbolic cone in three dimensions accumulates as a planar ellipse or hyperbola respectively. In plasticity theory, this approach delivers the yield criterion (see Collins and Houlsby, 1997, §§4 and 5, for example for soils and other geomaterials). The variables there are components of second order tensors, such as stress or strain-rate.

We have given some examples in (5), (12) and (17) of definitions of duality which can be transformed to Legendre duality by separated mappings (7) so that in each case the two types of duality co-exist and, in the sense of Theorem 1, are equivalent.

A fundamentally different type of duality is provided by the example of (1) in which

$$F(x_i, y, X_i, Y) = x_i X_i - yY \ . \tag{43}$$

This is bilinear, but it is distinct from Legendre duality defined using (4), and the following result comes from looking at (43) as it stands.

Theorem 6

When the hypotheses of Theorems 2 and 3 apply to $F = 0$ with (43),

$$x_i X_i = yY \tag{44}$$

implies

$$X_i = Y \frac{\partial y}{\partial x_i} \quad , \quad x_i = y \frac{\partial Y}{\partial X_i} \tag{45}$$

in terms of the assumed functions $y(x_i)$ and $Y(X_i)$.

The proof is immediate. A feature of (45) is that the gradient mappings of (25) and (27) are recovered on the *particular* sections

$$y(x_i) = 1 \quad \text{and} \quad Y(X_i) = 1 \,. \tag{46}$$

This is significant if both $y(x_i)$ and $Y(X_i)$ can be homogeneous functions of degree one, in which case those sections represent the essential shape of the functions. That fact is exploited by Hill (1987) to reach a pair of gradient mappings relating stress and strain-rate in a rigid/plastic solid, from starting assumptions different to our (44), which he reaches as a deduction. This applies, for example, to metals when the elastic deformation is negligible.

We have already shown in Theorem 5 that Legendre dual functions cannot both be homogeneous of degree one, and of course (46) does not define Legendre duality. The following result replaces Theorem 5.

Theorem 7

(a) When (44) and $(45)_1$ apply with $y \neq 0$, and $y(x_i)$ is homogeneous of degree r, then $r = 1$ or $Y = 0$.

(b) When $(45)_2$ also applies with $Y \neq 0$, then $Y(X_i)$ and $y(x_i)$ are both homogeneous of degree one.

Proof

(a) From (36), $(45)_1$ and (44) we deduce $(r-1)yY = 0$.

(b) From $(45)_2$ and (44) we get $\left(Y - X_i \dfrac{\partial Y}{\partial X_i}\right) y = 0$.

□

This establishes that when both the functions $y(x_i)$ and $Y(X_i)$ in (45) are non-zero and homogeneous, that homogeneity must be of order one.

Theorems 5(b), 6 and 7 make it clear that Legendre duality and the duality defined by (44) cannot be equivalent. It is easy to show, however, that they can co-exist in the sense described in terms of (2) and (3). Simple trial and error shows that

$$x_i X_i - yY = w_i W_i - p - P \tag{47}$$

can be satisfied by invertible transformations with $x_i = w_i$, $X_i = W_i$

and

$$p = \tfrac{1}{4}(y+Y)^2 \quad , \quad P = -\tfrac{1}{4}(y-Y)^2 \tag{48}$$

so that

$$y = \sqrt{p} + \sqrt{-P} \quad , \quad Y = \sqrt{p} - \sqrt{-P} \ . \tag{49}$$

Co-existence then occurs in the region $p \geq 0, P \leq 0$ and for all y and Y. Another option is to take $x_i = w_i, X_i = W_i, y = p \neq 0$ and

$$P = y(Y-1) \quad \text{so that} \quad Y = 1 + \frac{P}{p} \ . \tag{50}$$

However, these are not separated transformations of the type (7), but mixed transformations of the more general type (2). The clear-cut Theorem 1, which identifies the equivalence of two *functions* as in (8), does not apply. A given pole $x_i = a_i, y = b$ and its associated G-polar $a_i X_i = bY$ induce neither a unique transformed pole nor a unique transformed G-polar. Only the *values* of the two functions corresponding to (8) are the same, but not their functional forms, which become ambiguous because of the non-uniqueness just mentioned.

A particular feature of Legendre duality is conveyed by the following result, in an $n+1$ dimensional space spanned here by cartesian coordinates X_i, Y.

Theorem 8

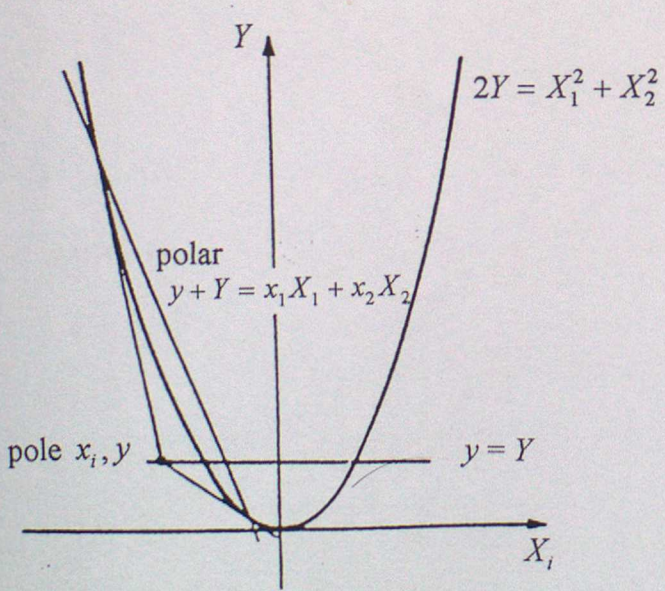
Suppose the pole x_i, y lies outside the paraboloid $2Y = X_i^2$. Construct a cone whose vertex is at the pole and whose generators touch the paraboloid. They do so in a plane section where the cone is intercepted by the polar $x_i X_i = y + Y$.

Proof

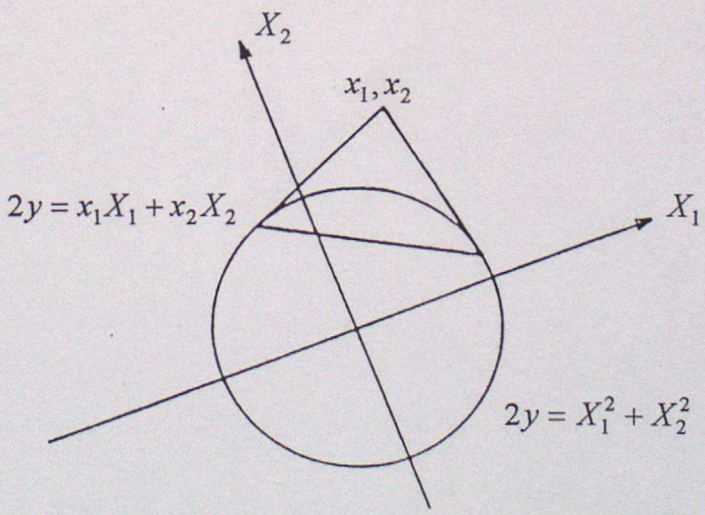
The vector which is perpendicular to the paraboloid at the point X_i, Y on it has components proportional to $-X_i, 1$. The join of this point to the pole has components $X_i - x_i, Y - y$. Because $2Y = X_i^2$, the scalar product of the two vectors is $-X_i(X_i - x_i) + Y - y = X_i x_i - Y - y$. This is zero if and only if the considered point also lies on the polar.

□

This classical result is illustrated in Fig. 6(a). A more special feature of duality, which appears in many geometry books and has been traced back in antiquity to Appolonius,



(a)



Section $y = Y$ through paraboloid, polar and pole.

(b)

Fig. 6 Legendre duality associated with (a) paraboloid and (b) circle.

emerges when attention is confined to the plane section $Y = y$ of Fig. 6(a) through the pole itself, as shown in Fig. 6(b). Then we see (for $n = 2$) the circle of Appolonius as the cross-section of the paraboloid, and the tangents from the pole meet the circle where the polar cuts it. This special case, when movement of the pole is confined to the plane $y = Y$, and the corresponding envelope of the polar in that plane is the dual curve, may be called Legendre duality with respect to a circle or sphere or hypersphere. Such specialization is sufficient for

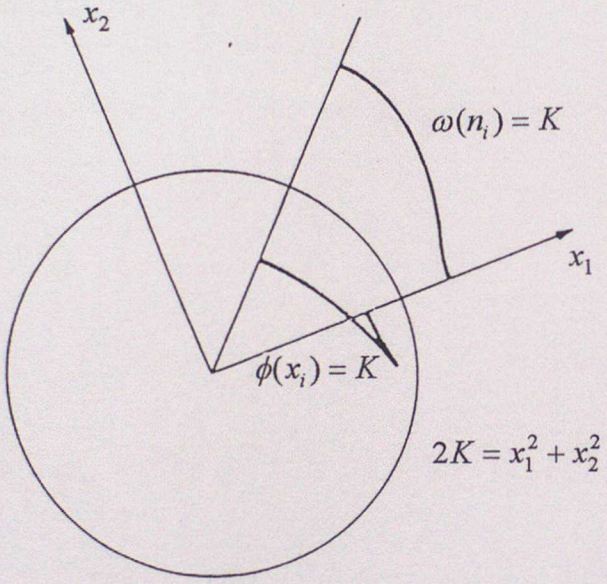


Fig. 7 Dual phase and slowness surfaces with respect to a circle.

some physical applications. Fig. 7 shows an example (Sewell 1982) from the propagation in a medium spanned by x_i of a plane wave whose normal is n_i with phase $\phi = n_i x_i - \omega t$ and wave speed $\omega \Big/ \left(n_i^2\right)^{1/2}$. At $t = 1$ we have $\phi + \omega = n_i x_i$, the starting point for Legendre duality. Dual sections $\phi(x_i) = K$ of a 'wave surface' and $\omega(n_i) = K$ of a 'slowness surface' emerge with respect to a circle $x_1^2 = 2K$. They may possess a corresponding cusp and inflexion as shown.

The construction of meshes or grids for efficient numerical work is of fundamental importance in applied mathematics. Meshes can occur in pairs, linked by a definition of duality

for meshes which is, at first sight, quite distinct from definitions of duality which start from statements like equation (1).

For example, a *plane mesh* is defined to be a set of polygons and their associated connected edges. Each edge is either closed, by two corners, or half-open extending to infinity at one end and closed by one corner at the other.

Two meshes A and B can be said to be *dual to each other* if two correspondences, each between the corners of one and the polygons of the other, can be set up so that the following three properties hold:

(a) the number of corners in A is the same as the number of polygons in B ;

(b) the number of corners in B is the same as the number of polygons in A ;

(c) each pair of corners of A that correspond to adjacent polygons of B is joined by an edge which can be put into correspondence with the common edge of the associated polygons of B, and vice versa.

Chynoweth and Sewell (1990) compared mesh duality with Legendre duality, and established a number of connections, including the following result.

Theorem 9

A general pair of meshes, which satisfy the definition of duality for meshes, is the projection onto the plane of a pair of piecewise linear functions in three dimensions that are Legendre dual versions of each other.

The proof is in the paper cited, where it is also shown how those piecewise linear Legendre dual functions can be the tangent plane and chordal plane discretizations of a pair of smooth Legendre dual functions.

Additional properties often possessed by dual meshes are that they are (d) orthogonal and (e) convex. Two dual meshes are orthogonal if, when superimposed, the members of each

pair of corresponding edges are orthogonal. A convex mesh is such that every polygon in it is convex.

A simple example of Theorem 9 is provided by the well known Delauney mesh (1934) and Voronoi tessellation (1908). Four points in the plane in general lie neither on a circle nor a straight line, so there are four circumcircles through them taken three at a time. Two will enclose the fourth point and two will not. The latter case is called a Delaunay triangulation, and the construction can be continued through a mesh.

A Voronoi tessellation of the same set of points divides the plane into a number of polygons, each containing one such point, and consisting of those parts of the plane that are closer to that point than any of the other points. The edges of these polygons lie along the perpendicular bisectors of the sides of the Delaunay triangles. Hence the corners of the Voronoi polygons are the Delaunay circumcentres. Fig. 8 shows an example of a Delaunay

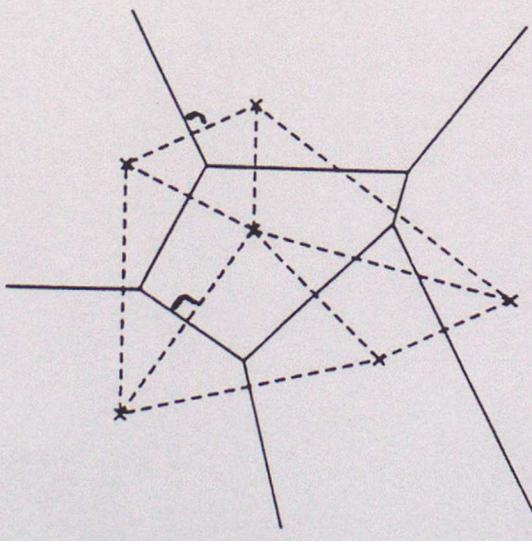


Fig. 8 Superimposed Delaunay triangulation and Voronoi tessellation.

triangulation of six points, and the associated Voronoi tessellation. The following result is proved by Chynoweth and Sewell (1990).

Theorem 10

The Delaunay triangulation and the Voronoi tessellation are that particular example of a pair of dual meshes which are the projections of tangent plane and chordal approximations to the same *paraboloid* $\frac{1}{2}X_i^2$ or $\frac{1}{2}x_i^2$.

A physical example of a Voronoi tessellation has been claimed in cosmology (Coles, 1990), where the galaxies are said to be distributed along the edges or "megawalls" of a particular Voronoi tessellation.

The definition of duality which we gave at the beginning of this Section began by setting to zero a function of variables which span two spaces. Such a function can be set down arbitrarily, and it need be neither continuous nor smooth in the first instance, or it may arise naturally from other considerations. In (1) it was a real function of $\mathbb{R}^{n+1} \times \mathbb{R}^{n+1}$, but other types of duality might be similarly defined using other types of functions (for example, complex ones) of variables spanning other types of spaces (again, for example, complex ones). The general possibilities may be unexplored.

3. Catastrophes

The presence in Legendre transformations of isolated singularities, like those represented in Figs. 4, 5 and 7 by cusps and inflexions which are dual to each other, is an indication of the ubiquitous role in science of the elementary catastrophes. There are many styles and contexts in which these can be introduced. Here we describe a straightforward one.

Suppose a scalar function $V(q_i, c_a)$ of n configuration variables q_i $(i = 1, \ldots, n)$ and m assignable parameters $c_\alpha (\alpha = 1, \ldots, m)$ is given. If it represents a *potential energy*, then the n equations

$$\frac{\partial V}{\partial q_i} = 0 \tag{51}$$

can be regarded as defining an *equilibrium surface* (Sewell, 1966) lying in the $n+m$ dimensional configuration/parameter space. In the mechanics of solids, for example, the q_i can be position coordinates, and the c_α can be 'control' parameters representing the values of loads, dimensions, moduli or imperfections.

The *bifurcation set* in parameter space is defined by eliminating the q_i between (51) and the determinantal equation

$$\left| \frac{\partial^2 V}{\partial q_i \, \partial q_j} \right| = 0 \, . \tag{52}$$

Points satisfying (52) on the equilibrium surface can be called a *stability boundary* there, because it consists of points where the stability of the equilibrium comes into question. At some or all of these points the function $V(q_i)$, for certain choices of the c_α in some range of test values, can cease to be a local minimum there.

The bifurcation set is the *projection* onto the parameter space of the stability boundary in configuration/parameter space. It is the bifurcation set which is the subject of deep theorems in catastrophe theory (e.g. see Poston and Stewart, 1978), where $V(q_i, c_\alpha)$ does not have to represent a potential energy in every applied context.

The list of *elementary catastrophes* includes the following.

Cuspoids

fold	$V(q,a) = q^3 + aq$
cusp	$V(q,a,b) = q^4 + bq^2 + aq$
swallowtail	$V(q,a,b,c) = q^5 + cq^3 + bq^2 + aq$
butterfly	$V(q,a,b,c,d) = q^6 + dq^4 + cq^3 + bq^2 + aq$

Umbilics

elliptic	$V(q,r,a,b,c) = q^2 r - r^3 + cq^2 + br + aq$

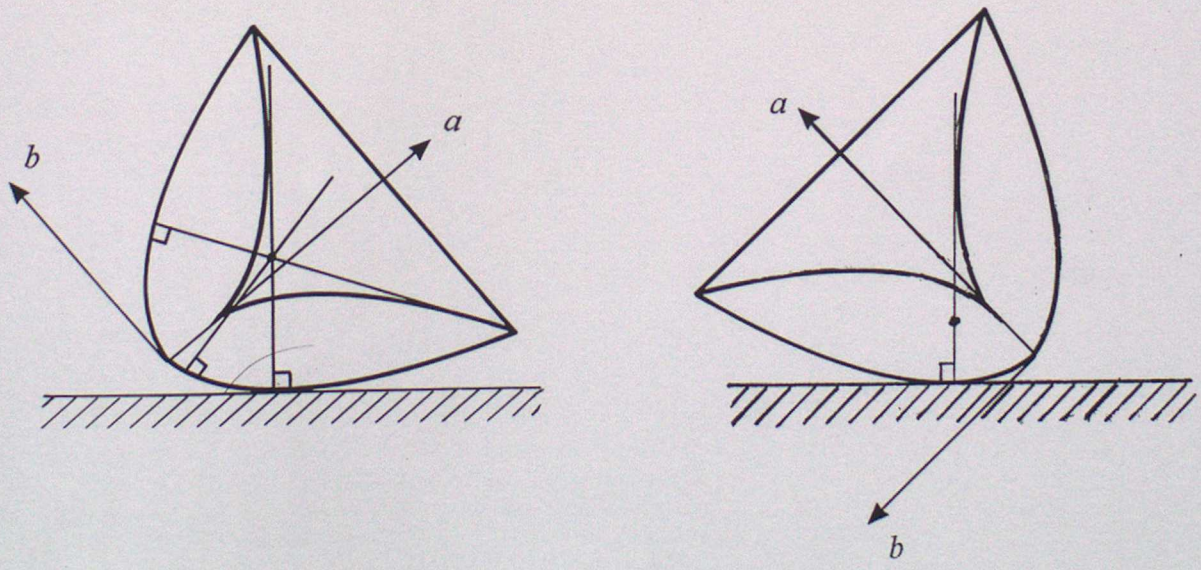
hyperbolic $V(q,r,a,b,c) = q^2r + r^3 + cq^2 + br + aq$

parabolic $V(q,r,a,b,c,d) = q^2r + r^4 + dr^2 + cq^2 + br + aq$

The following theorem summarizes a simple classroom demonstration (Sewell, 1978a, 1990) of an equilibrium surface with its stability boundary and bifurcation set, obtained by interpreting the cusp catastrophe quartic as a potential energy.

Theorem 11 (Parabolic catastrophe machine)

A mechanical system consists of a point mass attached to a lamina which occupies the inside of a parabola, and which is confined to a single vertical plane (Fig. 9). The system can either stand or roll on a horizontal plane, along the parabolic boundary. The envelope of the normals to the parabola is the bifurcation set of the cusp catastrophe quartic, and the latter is the effective potential energy. Quasi-static transit of the mass across the bifurcation set, from inside to out, induces a dynamic jump of the system to a distant equilibrium configuration.



(a) Three equilibria (b) One equilibrium

Fig. 9 Parabolic catastrophe machine.

Proof

With respect to cartesian a, b axes attached to the lamina, let the parabolic boundary $b^2 = 4a$ be described parametrically as $a = q^2, b = 2q$. When the mass is attached to an *interior* point with coordinates a, b, and when the system is in equilibrium with $q^2, 2q$ as the coordinates of that *boundary* point which is in contact with the plane, the equilibrium equation is $(b-2q)/(a-q^2) = -q$. This is because the mass must lie directly above the contact point, and therefore on the normal through it. The single parameter q specifies the configuration, and the adjustable coordinates a, b of the mass are the only control variables in this problem. The equilibrium equation can be rewritten

$$q^3 + q(2-a) - b = 0 \quad \text{or} \quad \partial V/\partial q = 0 \tag{53}$$

where the quartic cusp catastrophe

$$V(q,a,b) = \tfrac{1}{4}q^4 + \tfrac{1}{2}q^2(2-a) - bq \tag{54}$$

is evidently the effective potential energy (although not the actual one, but the first two q-derivatives of each can be shown to vanish together). Equation (53) defines the well-known folded cusp catastrophe equilibrium surface in q, a, b space, whose stability boundary is the locus of points on it where

$$\partial^2 V/\partial q^2 = 0 \quad \text{or} \quad a - 2 = 3q^2 \ . \tag{55}$$

Cross-sections of the surface with planes $a =$ constant for $a > 2$ are cubic curves in the q, b plane, having turning points at the stability boundary. The bifurcation set in the a, b plane is the envelope of the normals to the parabola, and this is obtained by eliminating q between (53) and (55) to obtain the cusped curve

$$27b^2 = 4(a-2)^3 \tag{56}$$

shown in Fig. 9. Thus the bifurcation set is tangible because it lies in the physical plane of the experiment.

The stability boundary is a single curve on the equilibrium surface. This boundary

encloses only points of instability 'inside' it, and possesses only points of stability outside it. When the mass is moved across the bifurcation set from 'inside' to out, the current stable equilibrium at an intersection of three normals (Fig. 9(a)) is lost, and a jump must occur, manifested by a roll of the parabola before equilibrium can be resumed, associated with what is then only one normal through the new control point (Fig. 9(b)). □

A string model of the equilibrium surface (53) can be made in three-dimensional q, a, b space, because (53) is a straight line in a, b space for each fixed q (Sewell, 1987, Fig. 2.8). Zeeman (1977) described many other physical applications of the elementary catastrophes.

The interest which the elementary catastrophes have from the viewpoint of transformation theory is that their bifurcation sets have already been determined, and there is a close relation between these (often multi-valued) functions and the Legendre duals (often having isolated singularities) of the elementary catastrophes. One such result is the following.

Theorem 12 (Ladder for the cuspoids)

The Legendre dual of the cuspoid polynomial of degree p, lying in a $(p-1)$ dimensional space, has the same shape as the bifurcation set of the cuspoid of degree $p+1$.

The proof of this has been given by Sewell (1977, 1987, Theorem 2.7) in two different ways. The result is graphically expressed in Fig. 10.

The lowest link, and the first two steps on the ladder, have already been illustrated in Figs. 5 and 9. Fig. 5 shows $q^3 + aq$ and its cusped Legendre dual, and the latter curve also emerges in Fig. 9 as the bifurcation set in parameter space of the quartic. The second link, connecting the third and fourth steps on the ladder, does not have a simple algebraic expression, but Theorem 12 allows us to see that the Legendre dual of $Y(X) = \frac{1}{4}X^4 + \frac{1}{2}bX^2$ for each different fixed choice of b has the swallowtail shape shown in Fig. 11(b) (from

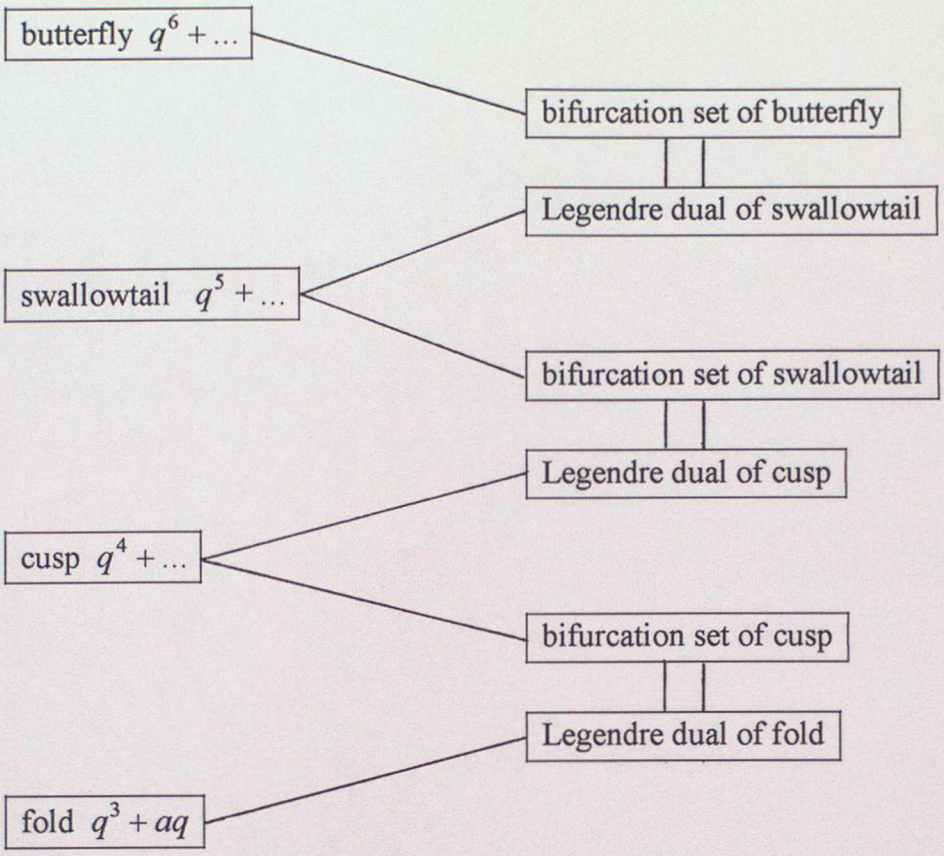
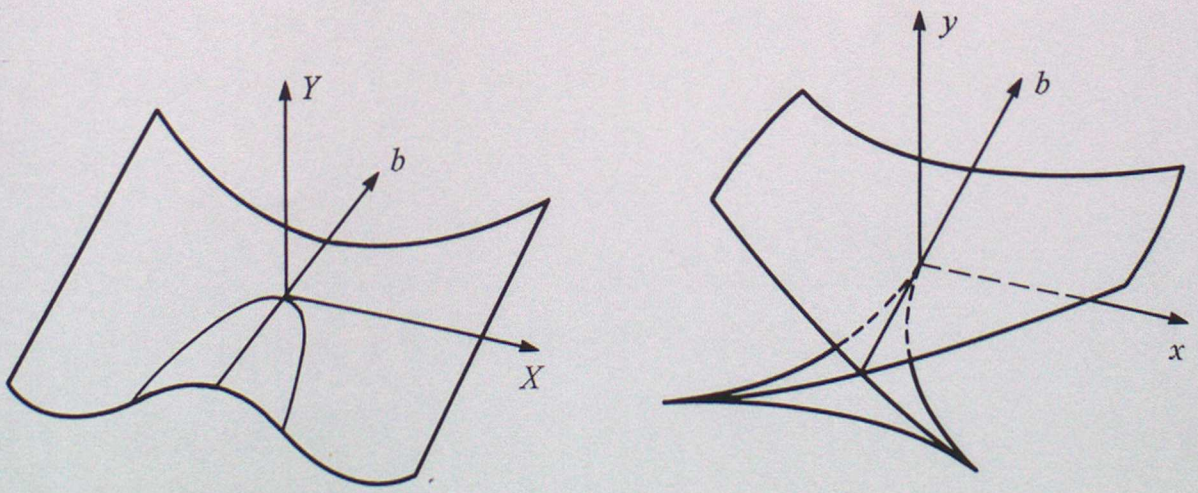


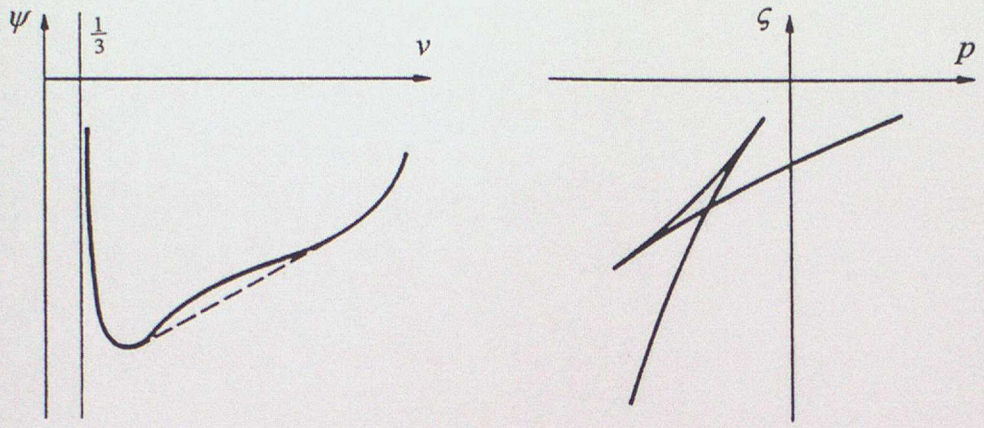
Fig. 10 Ladder for the cuspoids

Sewell, 1987, Fig. 2.11). Here we use the notation of (25) and (27). A physical example from thermodynamics is shown in Fig. 12, for a cross-section diffeomorphic to one for $b<0$ in Fig. 11.



(a) $Y=\frac{1}{4}X^4+\frac{1}{2}bX^2$ (b) Dual $y=y(x,b)$ of (a)

Fig. 11 Cusp potential and its swallowtail dual.



(a) Free energy $\psi(v)$. (b) Swallowtail free enthalpy $\zeta(p)$

Fig. 12 Dual van der Waals potentials for constant temperature.

The umbilics provide a helpful illustration of how to calculate Legendre transformations involving dual isolated singularities when two active variables are present. For example, the hyperbolic umbilic with $c<0$ and $b=a=0$, i.e. $X_1^2X_2+X_2^3+cX_1^2$, has

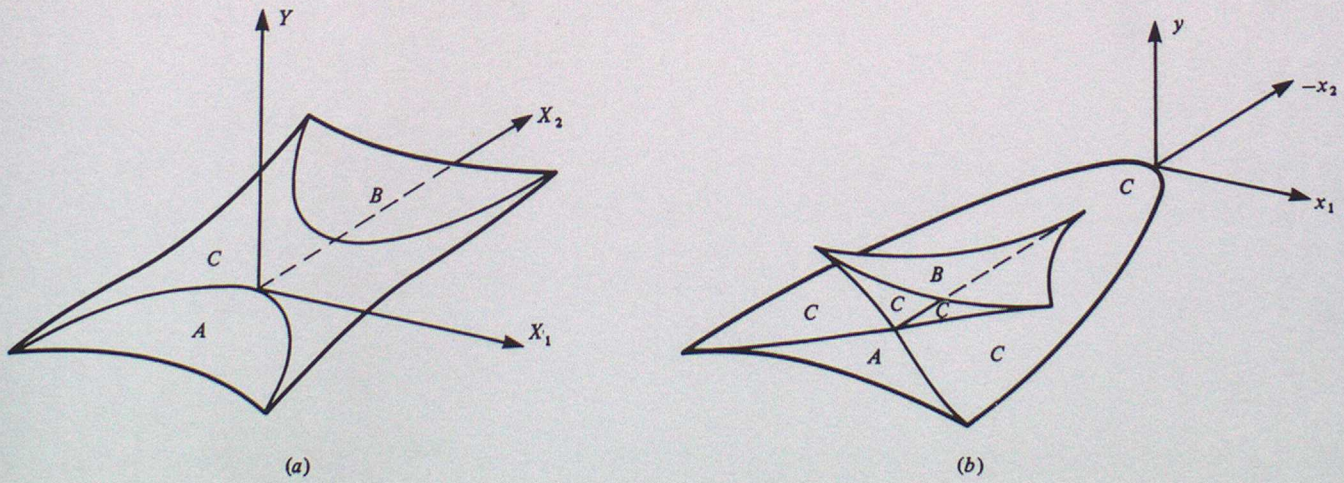


Fig. 13(a) Hyperbolic umbilic and (b) its Legendre dual

the single-valued shape shown in Fig. 13(a), and it can be shown (Sewell, 1978b, 1987) that its Legendre dual is the surface shown in Fig. 13(b). The latter has 0, 2 or 4 valued parts separated by cusped or swallowtail singularities. The parts labelled A, B and C are respectively concave, convex and saddle-shaped in each diagram.

Other styles of approach to the study of Legendrian singularities are to be found in the work of Arnold (e.g. 1989).

Convexification of Legendre dual functions can be significant for physical reasons. This can be achieved in Fig.12, for example, by replacing the 'hump' in Fig. 12(a) by the straight dotted line, and the Legendre dual of the result is obtained by deleting the triangular tail from the swallowtail in Fig. 12(b). This leaves a concave function having an isolated jump in gradient. The same can be done for each value of b in Fig. 11.

Atmospheric fronts are isolated locations where there are jumps in temperature and wind velocity. A preliminary exploration of how part of the parabolic umbilic and its Legendre dual could be convexified was carried out by Chynoweth, Porter and Sewell (1988). They showed how a front might be regarded as a self-intersection line terminating at a

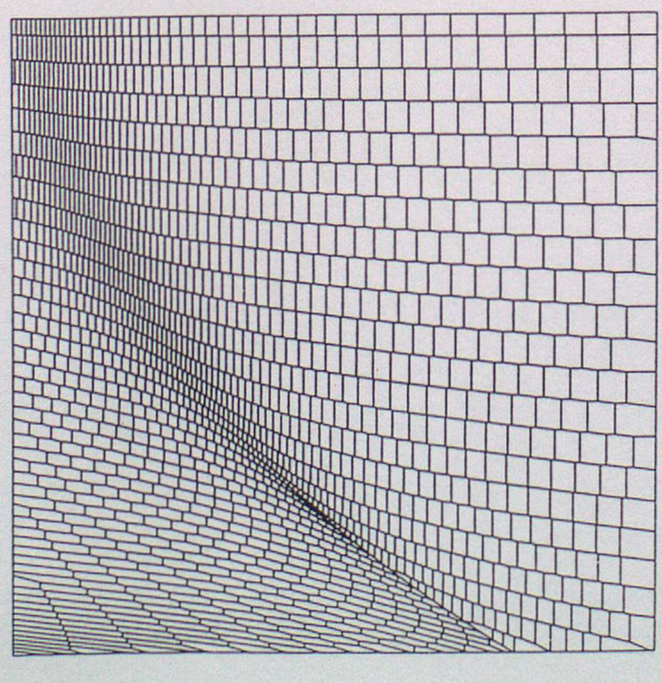


Fig. 14 Discretization of an atmospheric front, in a vertical plane.

swallowtail point on a Legendre dual surface. A discretization of such a model is shown in Fig. 14, which is intended to represent a single vertical plane in the atmosphere, with the ground at the bottom.

4. Upper and lower bounds

Suppose that $F[x,u]$ is a functional defined on a pair of inner product spaces. The inner products may, in general, be different and we denote them by (x_1,x_2) and $\langle u_1,u_2\rangle$. Suppose that $F[x,u]$ is a *saddle functional* in the sense that

$$F_1 - F_2 - \left(x_1 - x_2, \left.\frac{\delta F}{\delta x}\right|_1\right) - \left\langle u_1 - u_2, \left.\frac{\delta F}{\delta u}\right|_2\right\rangle \geq 0 \tag{57}$$

for all pairs x_1,u_1 and x_2,u_2 in the domain of $F[x,u]$, where the partial functional derivatives denoted using δ are evaluated at the indicated points.

Theorem 13

Any solution x_0,u_0 of the problem defined by

$$\begin{aligned} &\frac{\delta F}{\delta x} \leq 0 \quad (\alpha) && \frac{\delta F}{\delta u} \geq 0 \quad (\beta) \\ &x \geq 0 \quad (\beta) && u \geq 0 \quad (\alpha) \\ &\left(x, \frac{\delta F}{\delta x}\right) = 0 && \left\langle u, \frac{\delta F}{\delta u}\right\rangle = 0 \end{aligned} \tag{58}$$

is characterized by the following upper and lower bounds

$$F_\alpha - \left(x_\alpha, \left.\frac{\delta F}{\delta x}\right|_\alpha\right) \geq F[x_0,u_0] \geq F_\beta - \left\langle u_\beta, \left.\frac{\delta F}{\delta u}\right|_\beta\right\rangle . \tag{59}$$

Here an α subscript specifies evaluation at any point which is required to satisfy only the conditions (58α), and similarly β.

If the left trio of (58) is replaced by the left of

$$\frac{\delta F}{\delta x} = 0 \quad (\alpha) \quad , \quad \frac{\delta F}{\delta u} = 0 \quad (\beta) \ , \tag{60}$$

or the right trio of (58) is replaced by the right of (60), or if both replacements are made, then the corresponding simplification of (59) applies.

Proof

Choose $x_1 = x_\alpha,\ u_1 = u_\alpha,\ x_2 = x_0,\ u_2 = u_0$ in (57). This delivers the upper bound.

Choose $x_1 = x_0,\ u_1 = u_0,\ x_2 = x_\beta,\ u_2 = u_\beta$ in (57). This delivers the lower bound.

□

Numerous examples of this result, in a slightly different notation, are detailed by Sewell (1987), together with sufficient conditions for (57) to be satisfied. Theorem 13 applies, in particular, to the very common subclass of problems in which there is a pair of linear operators T and T^* which are adjoint to each other in the sense that

$$\left(x, T^* u\right) = \left\langle u, Tx \right\rangle \tag{61}$$

for all x and u in the domains of the respective operators, and in which

$$F[x,u] = (x, T^* u) - H[x,u] \ , \tag{62}$$

where $H[x, u]$ is another functional. This is used in (60) to produce (Noble, 1964) the generalized hamiltonian problem

$$T^* u = \frac{\delta H}{\delta x} \quad (\alpha) \quad , \quad Tx = \frac{\delta H}{\delta u} \quad (\beta) \ . \tag{63}$$

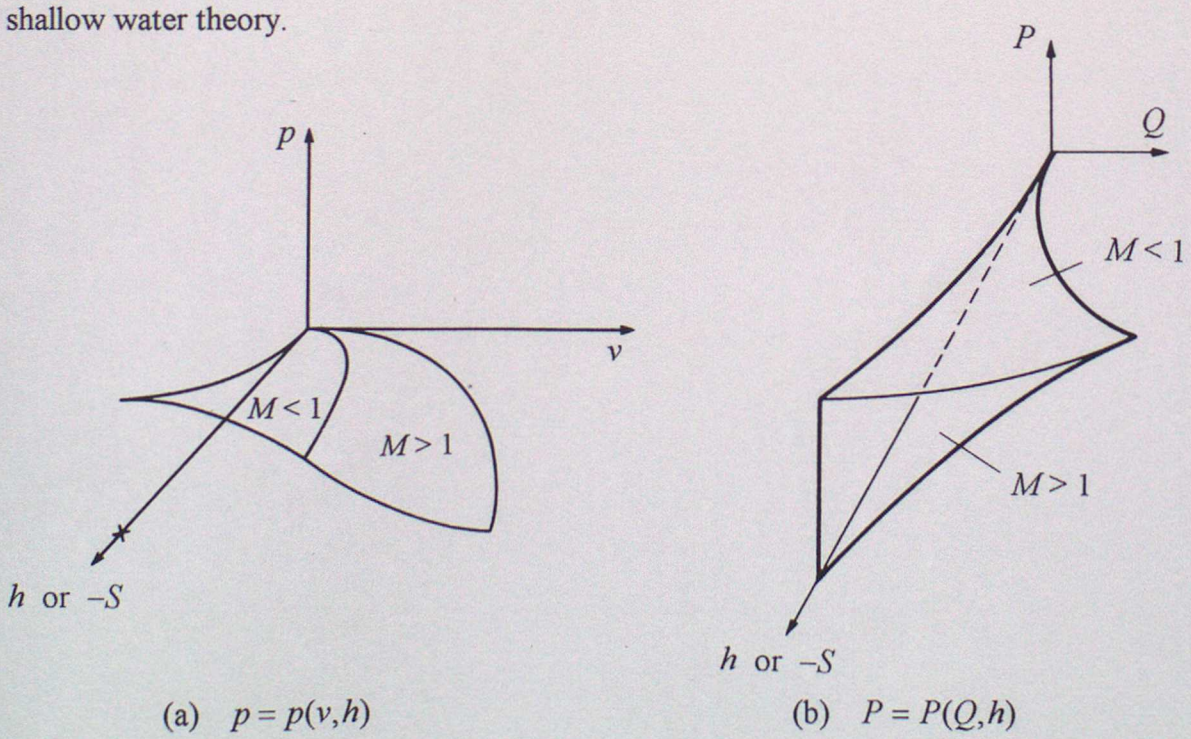
The true adjointness property (61) includes integration by parts, for example, in one or more dimensions, with explicit incorporation of the boundary terms, and also integral operators.

We cite the following example in gas dynamics where parts of surfaces diffeomorphic to those illustrated in Fig. 11, for the cusp potential and its swallowtail dual, play a central role in satisfying the hypothesis (57).

Associated with each particle of perfect fluid, having any equation of state and in any motion, there is a function $p(v, h, S)$ expressing the pressure p in terms of speed v, total energy $h = H + \frac{1}{2}v^2$ where H is enthalpy, and entropy S, such that from thermodynamics alone

$$\frac{\partial p}{\partial v} = -\rho v \quad , \quad \frac{\partial^2 p}{\partial v^2} = \rho\left(\frac{v^2}{c^2} - 1\right) \quad , \quad \frac{\partial^3 p}{\partial v^3} \neq 0 \tag{64}$$

where ρ is density and c is the sound velocity. Defining mass flux $Q = \rho v$ and flow stress $P = Qv + p$ allows us to construct a Legendre dual function $P(Q,h,S)$ with $\partial P/\partial Q = v$ in any motion (Sewell, 1963). Geometrically, $p(v, h, S)$ and $P(Q, h, S)$ may be described as *constitutive surfaces* whose shape can be determined in advance for any motion, and which were shown by Sewell and Porter (1980) to have the form illustrated in Fig. 15(a) and (b), for each fixed h or for each fixed S (with $M = v/c$). Similar, but simpler, results apply for shallow water theory.



(a) $p = p(v,h)$ (b) $P = P(Q,h)$

Fig. 15 Legendre dual constitutive surfaces in any gas flow.

Introducing the vector velocity $\boldsymbol{v}$ and mass flow $\mathbf{Q} = \rho \boldsymbol{v}$, the Legendre transformation between $P(\mathbf{Q})$ and $-p(\boldsymbol{v})$ for each fixed h and S can be rephrased as

$$\frac{\partial p}{\partial \boldsymbol{v}} = -\mathbf{Q} \quad , \quad \frac{\partial P}{\partial \mathbf{Q}} = \boldsymbol{v} \quad , \quad P = p + \mathbf{Q}.\boldsymbol{v} \ . \tag{65}$$

Restricting attention now to homentropic ($S \equiv$ constant), homenergic ($h \equiv$ constant) flow in a closed region V whose boundary $A + B$ has two parts, with outward normal $\mathbf{n}$, this flow is irrotational and is governed by the equations

$$\boxed{\begin{aligned} \nabla\phi &= \frac{\partial P}{\partial \mathbf{Q}} \ \text{ in } V \\ -\mathbf{n}\phi &= -\mathbf{n}a \ \text{ on } A \end{aligned}} \quad (\alpha) \qquad \boxed{\begin{aligned} -\nabla.\mathbf{Q} &= 0 \ \text{ in } V \\ \mathbf{n}.\mathbf{Q} &= b \ \text{ on } B \end{aligned}} \quad (\beta) \tag{66}$$

Here ϕ is the velocity potential, and a and b are assigned functions of position on the two parts A and B of the boundary respectively. Equations (66) have the structure (63), and the volume and surface components of each left side define the composite operators T and T^*, by comparing (61) with Green's theorem in the form

$$\int \mathbf{Q}.\nabla\phi \, dV - \int \mathbf{Q}.\mathbf{n}\phi \, dA = -\int \phi \nabla.\mathbf{Q} \, dV + \int \phi \mathbf{n}.\mathbf{Q} \, dB \ . \tag{67}$$

Further details are described by Sewell (1987, §5.3(v)), where it is shown that solutions of (66) make the functional

$$F[\mathbf{Q},\phi] = \int (\mathbf{Q}.\nabla\phi - P(\mathbf{Q}))dV - \int (\phi - a)\mathbf{n}.\mathbf{Q} \, dA - \int b\phi \, dB \tag{68}$$

stationary with respect to free variations of $\mathbf{Q}$ and ϕ.

Theorem 14

When the flow is also subsonic, the functions $P(\mathbf{Q})$ and $-p(\boldsymbol{v})$ are both convex, (68) is a saddle functional, and the upper and lower bounds in (59) are, respectively

$$-\int p(\nabla\phi)dV - \int b\phi \, dB \quad \text{for any } \phi = a \text{ on } A, \text{ and}$$

$$-\int P(\mathbf{Q})dV + \int a\,\mathbf{n}.\mathbf{Q}\,dA \quad \text{for any divergenceless } \mathbf{Q}$$

which satisfies $\mathbf{n}.\mathbf{Q} = b$.

The proof is an immediate consequence of Theorem 13, and the bounds can be minimized and maximized, respectively, to give a common value provided by a unique solution to the whole problem.

5. Lift and contact transformations

We begin this topic by motivating the terminology with a simple case. Using cartesian coordinates x, y, z in $\mathbb{R}^3$, the equations

$$x = x(u) \quad , \quad y = y(u) \quad , \quad z = z(u) \ , \tag{71}$$

with given functions of a scalar parameter u on the right, describe a space curve. We can say that it is a *lift* of the plane curve $x = x(u),\ y = y(u)$ if and only if

$$\frac{dy}{du} = z\frac{dx}{du} \ . \tag{72}$$

A *regular* point has $dx/du \neq 0$, and in its neighbourhood we could use x as the parameter in place of u, so that $z = dy/dx$. Then the height of the lift in $\mathbb{R}^3$ is the gradient of the plane curve in $\mathbb{R}^2$. This motivates the term 'lift'. Two different lifts which intersect at a regular point therefore have projections which are in *contact*, in the sense that they meet and have the same tangent line there. *Singular* points, where $dx/du = 0$, can also be handled with further hypotheses (Sewell and Roulstone, 1994, §2), and so we retain the symmetric form (71) with unspecified parameter u.

A *lift transformation* is a mapping

$$X = X(x,y,z) \quad , \quad Y = Y(x,y,z) \quad , \quad Z = Z(x,y,z) \tag{73}$$

of $\mathbb{R}^3$ to $\mathbb{R}^3$ which maps a lift (71) with (72) into another lift. The given functions on the

right of (73) therefore become new functions $X(u)$, $Y(u)$, $Z(u)$ with the property

$$\frac{dY}{du} = Z\frac{dX}{du} \quad . \tag{74}$$

A lift transformation may be either *general* or *special*, depending on whether or not it maps every lift, or only some particular lift, into another lift (Sewell and Roulstone, 1994, §3).

A *contact* transformation is a lift transformation applied to a *pair* of lifts which intersect in the first $\mathbb{R}^3$. It therefore transforms contact in the associated first $\mathbb{R}^2$ into contact in the second $\mathbb{R}^2$.

The following shorthand for combinations of the derivatives of (73) is convenient,

$$\alpha = \frac{\partial Y}{\partial x} - Z\frac{\partial X}{\partial x} \quad , \quad \beta = \frac{\partial Y}{\partial y} - Z\frac{\partial X}{\partial y} \quad , \quad \gamma = \frac{\partial Y}{\partial z} - Z\frac{\partial X}{\partial z} \quad . \tag{75}$$

Theorem 15

Necessary and sufficient conditions for (73) to have the property

$$\frac{dY}{du} - Z\frac{dX}{du} = \beta\left(\frac{dy}{du} - z\frac{dx}{du}\right) \tag{76}$$

for every smooth space curve (71) are that

$$\alpha + \beta z = 0 \quad \text{and} \quad \gamma = 0 \, . \tag{77}$$

Then

$$Z = \frac{\dfrac{\partial Y}{\partial x} + z\dfrac{\partial Y}{\partial y}}{\dfrac{\partial X}{\partial x} + z\dfrac{\partial X}{\partial y}} \quad \text{or} \quad Z = \frac{\dfrac{\partial Y}{\partial z}}{\dfrac{\partial X}{\partial z}} \tag{78}$$

or both, where $\partial X/\partial x + z\,\partial X/\partial y \neq 0$ or $\partial X/\partial z \neq 0$ or both, respectively. If

$$\beta \neq 0 \tag{79}$$

also, (73) transforms any lift into another lift.

A proof of this result is given *ab initio* by Sewell and Roulstone (1994, §3). In

particular we see from (76) that (72) and (79) imply (74). The subject is, of course, an old one, and other styles of exposition include those of Carathéodory (1982) and Arnold (1989). When a certain level of familiarity is reached, it is conventional to omit du from (76) so that the parameter u ceases to be mentioned. In some parts of the literature $dy - zdx$ would be called the Cartan form (cf. Rubtsov and Roulstone, 1997).

It is evident, for example from Theorem 2, that some smoothness hypotheses are desirable for some purposes. The next result shows how a particular choice of the function in (1) is induced when certain hypotheses in Theorem 15 are satisfied, in which case a lift transformation induces a definition of duality.

Theorem 16

Any transformation of the form

$$X = X(x,y,z) \quad , \quad Y = Y(x,y,z) \quad , \quad Z = \frac{\partial Y / \partial z}{\partial X / \partial z} \tag{80}$$

in which

$$\frac{\partial X}{\partial z} \neq 0 \tag{81}$$

permits the construction of a function $F(x,y,X,Y)$ with

$$\frac{\partial F}{\partial x} = \alpha \quad , \quad \frac{\partial F}{\partial y} = \beta \quad , \quad \frac{\partial F}{\partial X} = Z \quad , \quad \frac{\partial F}{\partial Y} = -1 \tag{82}$$

which can be used as the starting point for duality properties, for example in Theorem 2.

Proof

Hypothesis (81) allows the local construction of an inverse function $z(x,y,X)$ from $(73)_1$, and thence from $(73)_2$ a function

$$F(x,y,X,Y) = Y\left[x,y,z(x,y,X)\right] - Y \tag{83}$$

from which (82) follows by the chain rule. □

This is part of a result given by Sewell and Roulstone (1994, Theorem 10), where the duality implied by setting $F = 0$ is further explored, including the consequences of Theorem 2 for Legendre duality and semi-geostrophic duality, and the further consequences of requiring $(77)_1$ and (79) to hold so that (80) is a lift transformation.

There are many available examples of lift transformations (more conventionally called contact transformations, even when the context merely requires the lift of only one curve, rather than a pair of curves). Sewell and Roulstone (1994, §§4 to 7) show how to deduce some families of lift transformations, from $\mathbb{R}^3$ to $\mathbb{R}^3$ and also from $\mathbb{R}^{2n+1}$ to $\mathbb{R}^{2n+1}$, by starting from polynomial hypotheses.

In the semi-geostrophic theory of meteorology it turns out that a particular lift transformation from $\mathbb{R}^5$ to $\mathbb{R}^5$, and belonging to these polynomial families, has a special role (Sewell and Roulstone, 1994, Theorem 11; Roulstone and Sewell, 1997, Theorem 1). We use the notation of the latter paper to describe a simple form of it.

Theorem 17

The transformation

$$\left.\begin{aligned} X_i &= \frac{x_i}{\mu} + z_i \,, \\ \Psi &= \psi + \tfrac{1}{2}\mu z_i^2 \,, \\ Z_i &= \mu z_i \end{aligned}\right\} \qquad (84)$$

from x_i, ψ, z_i space to X_i, Ψ, Z_i space, for $i = 1, 2$ and constant $\mu \neq 0$, is a lift transformation.

When it is also hypothesized that ψ and z_i are functions

$$\psi = \psi(x_i, t) \quad , \quad z_i = \frac{\partial \psi}{\partial x_i} \tag{85}$$

then (84) contains the gradient starting point

$$X_i = \frac{\partial P}{\partial x_i} \quad , \quad P(x_i, t) = \tfrac{1}{2\mu} x_i^2 + \psi(x_i, t) \tag{86}$$

for a Legendre transformation with dual function $S(X_i, t)$ such that

$$x_i = \frac{\partial S}{\partial X_i} \quad , \quad P + S = x_i X_i \quad , \quad \frac{\partial P}{\partial t} + \frac{\partial S}{\partial t} = 0 \ . \tag{87}$$

This allows us to construct another function

$$\Psi(X_i, t) = \tfrac{1}{2} \mu X_i^2 - S(X_i, t) \tag{88}$$

such that $(84)_3$ implies

$$\frac{\partial \Psi}{\partial X_i} = Z_i \ . \tag{89}$$

Also

$$\frac{\partial \Psi}{\partial t} = \frac{\partial \psi}{\partial t} \ . \tag{90}$$

Proof

This is by straightforward manipulation (Roulstone and Sewell, 1996, Theorems 1, 2, 3, 4). It uses the intermediate fact that, when $(85)_1$ and $(86)_2$ are hypothesized with (84), but not yet $(85)_2$ or $(86)_1$, the values of

$$\Psi = P - x_i X_i + \tfrac{1}{2} \mu X_i^2 \ . \tag{91}$$

□

The result shows clearly the effects of introducing a sequence of hypotheses (84) and (85) to reach (88) and (89). Fig. 16 is a schematic representation of the sequence of hypotheses and their consequences in Theorem 17.

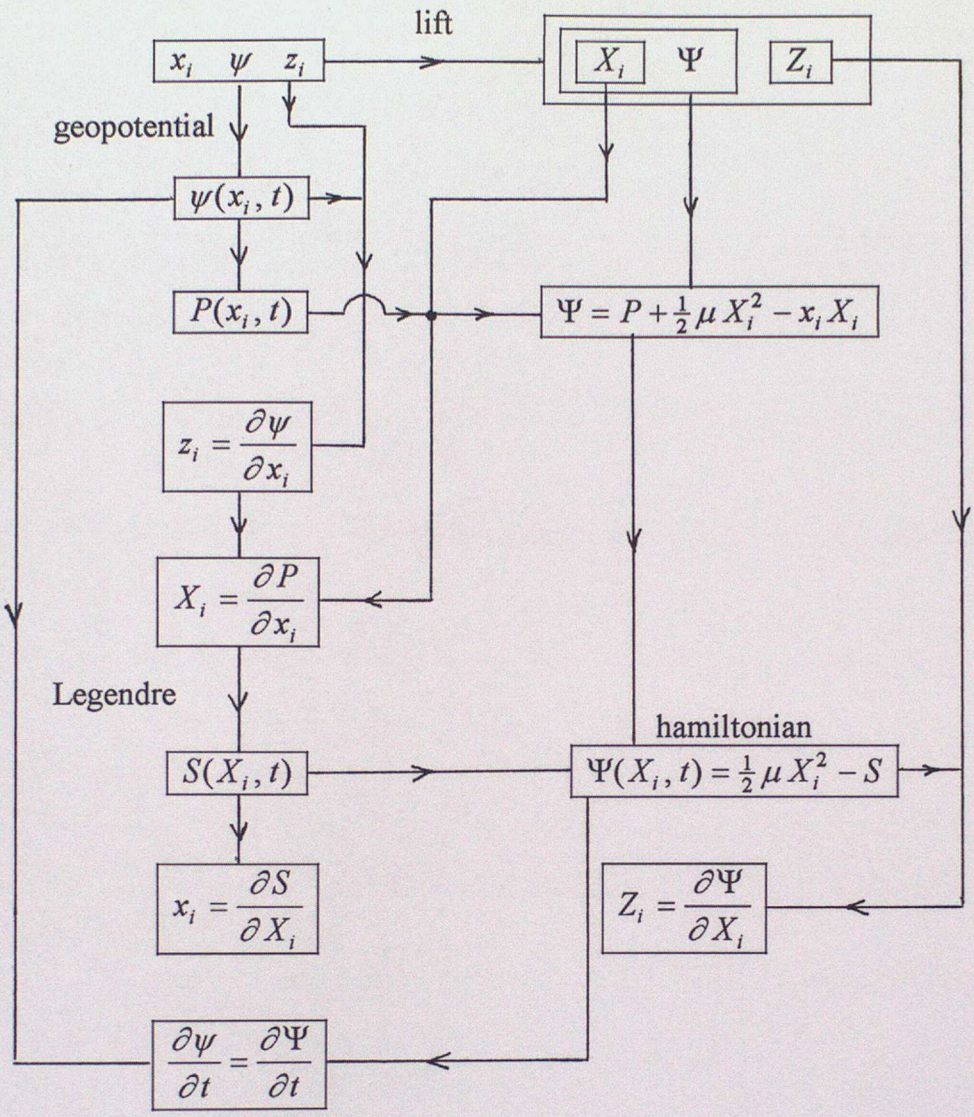


Fig. 16 Schematic version of Theorem 17.

In the semi-geostrophic theory of meteorology on the so-called f-plane we can put $\mu = 1$ and it turns out that, apart from scale factors, the x_i and ψ represent the horizontal coordinates, in physical space, and the geopotential respectively. To identify with (6) we replace the p, y there with P, ψ here. The Ψ here turns out to be the value of a hamiltonian. The X_i represent the geostrophic coordinates, and to further identify with (6)

we replace the P, Y there with the Ψ, S here. Then (91) is the same as putting $G = 0$ in (5) to generate what was called semi-geostrophic duality there.

6. Quartets of Legendre transformations

The structures in Theorem 17 are sufficient to encapsulate shallow water theory with constant Coriolis parameter, but the associated atmospheric theory requires incorporation of the vertical height and potential temperature. These require augmentation of Theorem 17 in a new direction. If x_α represents vertical height, hypothesis $(85)_1$ is replaced by

$$\psi = \psi(x_i, x_\alpha, t) \ . \tag{92}$$

Potential temperature is then represented by a new variable

$$X_\alpha = \frac{\partial P}{\partial x_\alpha}, \quad \text{where} \quad P(x_i, x_\alpha, t) = \frac{1}{2\mu} x_i^2 + \psi(x_i, x_\alpha, t) \tag{93}$$

is the new definition of the modified geopotential function. This leads to a quartet of Legendre transformations (Chynoweth and Sewell, 1989) summarized in Theorem 18 and Fig. 17.

Theorem 18

Figure 17 displays four alternative starting functions, in the boxes, from which to construct Legendre transformations whose active variables are written next to the links joining the boxes. The gradient properties of the quartet include

$$P + S = x_i X_i \quad , \quad X_i = \frac{\partial P}{\partial x_i} \quad , \quad x_i = \frac{\partial S}{\partial X_i} \ , \tag{94}$$

$$S - R = -x_\alpha X_\alpha \ , \quad X_\alpha = -\frac{\partial S}{\partial x_\alpha} \ , \quad x_\alpha = \frac{\partial R}{\partial X_\alpha} \ . \tag{95}$$

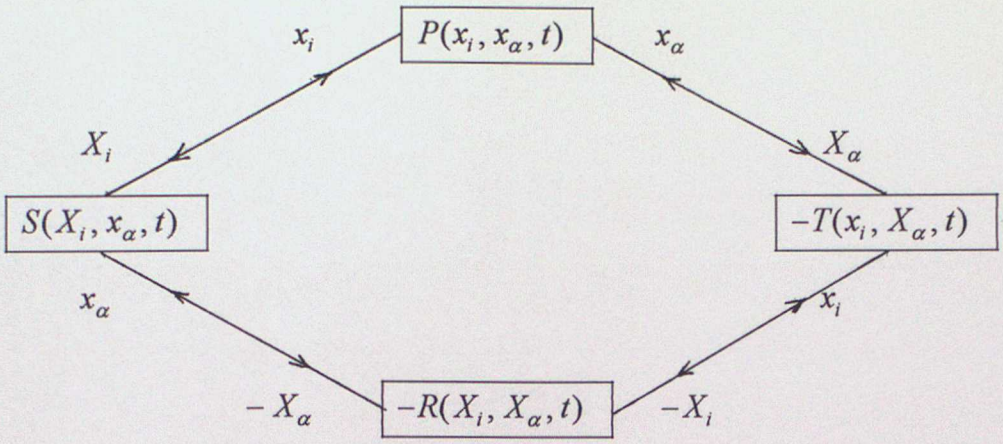


Fig. 17 Quartet of Legendre transformations.

The proof of such results is essentially by the chain rule, as explained by Sewell (1987, §§2.5 to 2.7) and Chynoweth and Sewell (1989, 1991), where it is also shown how convexity and concavity properties can be transmitted round the quartet of functions. A very simple example of such transmission is conveyed by the Morse functions in Fig. 18. This feature

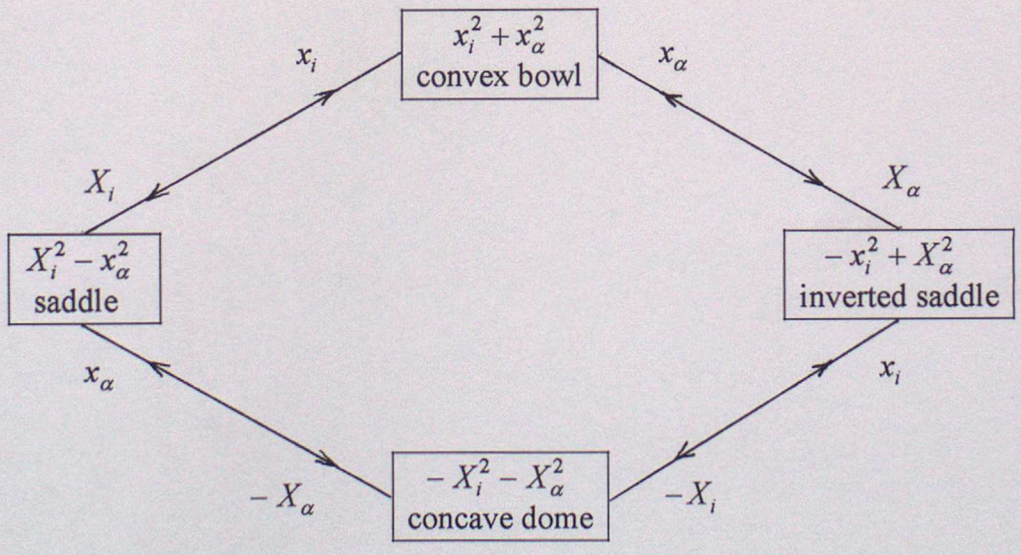


Fig. 18 Legendre chain of Morse functions.

applied to the general functions in Fig. 17 allows the deduction, at each t, of the synonymity property

$$\text{convex } P(x_i, x_\alpha) \;\Leftrightarrow\; \text{saddle } S(X_i, x_\alpha) \;\Leftrightarrow \text{convex } R(X_i, X_\alpha) \tag{96}$$

for corresponding parts of the indicated functions.

The interest of (96) in semi-geostrophic meteorology derives from the fact that, aside from a positive scale factor, the jacobians of the functions can be identified with the potential vorticity q (say) via

$$q = \frac{\partial^2 P}{\partial(x_i, x_\alpha)} = -\frac{\partial^2 S}{\partial x_\alpha^2}\left[\frac{\partial^2 S}{\partial(X_i, X_j)}\right]^{-1} = \left[\frac{\partial^2 R}{\partial(X_i, X_\alpha)}\right]^{-1} \tag{97}$$

(Chynoweth and Sewell, 1989, 1991).

When q is regarded as an *assigned* function $q(X_i, X_\alpha, t)$, the first and last of (97) comprise a Monge-Ampère equation for the determination of $R(X_i, X_\alpha, t)$. This approach is now a dominant theme in semi-geostrophic meteorology. From (97) we see that $q > 0$ where (96) holds at each t.

The property (96) is not a necessary feature of Fig. 17, but it may be the one of most physical interest. It might apply on only part of each corresponding function. Each Legendre transformation may also admit singularities, for example like the isolated singularities illustrated in Fig. 5, 7, 9, 11, 12, 13 and 14. The description of some aspects of catastrophe theory in §3 is motivated by the need to understand what types of singularity might appear in Fig. 17.

A relatively simple example which can portray the idea of an atmospheric front is generated by the cusp catastrophe potential

$$S(X_i, x_\alpha) = X_i^4 + x_\alpha X_i^2 \tag{98}$$

for $i = \alpha = 1$, expressed in so-called geostrophic coordinates. Chynoweth and Sewell (1989, Figs. 2 and 3) showed that (98) generates the quartet of dual functions shown in Fig. 19, and their convexified version is shown in Fig. 20. The latter exhibits a crease in the convexified $P(x_i, x_\alpha)$, which represents a front because the gradient discontinuities of $X_i = \partial P / \partial x_i$

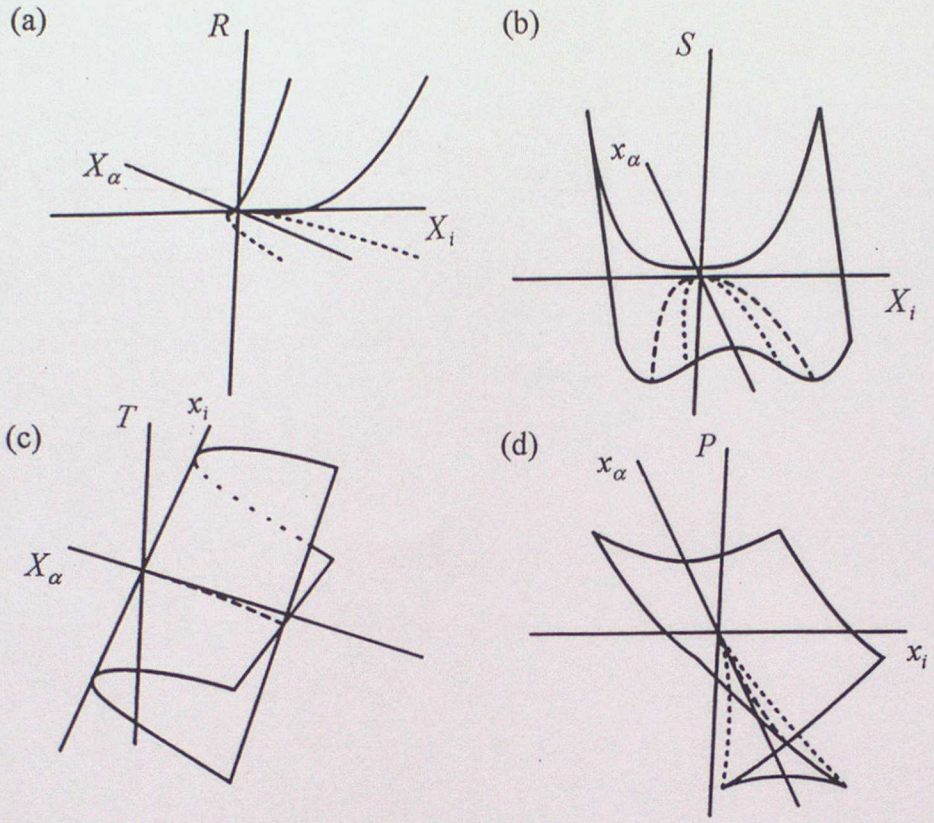


Fig. 19 Legendre quartet generated by $S(X_i, x_\alpha) = X_i^4 + x_\alpha X_i^2$.

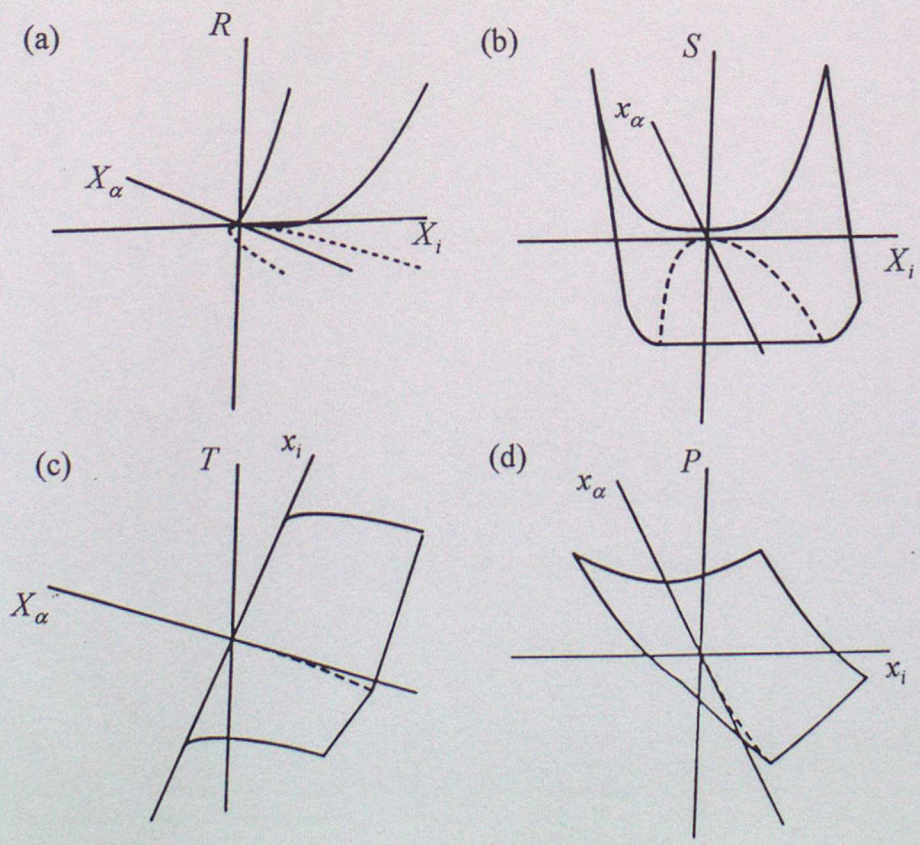


Fig. 20 Convexified version of Fig. 19.

and $X_\alpha = \partial P / \partial x_\alpha$ across the crease represent discontinuities of velocity and temperature. In Fig. 19(d) $P(x_i, x_\alpha)$ exhibits a swallowtail, thereby illustrating two of the steps in the ladder for the cuspoids in Fig. 10, because $P(x_i, x_\alpha)$ is the Legendre dual of (98) with active X_i and passive x_α. Cullen and Purser (1984) realized that, at each fixed t, convexity of the modified geopotential, here $P(x_i, x_\alpha)$, would be necessary for stability in semi-geostrophic applications. This particular example (98) is associated with a value of $q = 0$, as (97) shows. More details are given in §14 of Chynoweth and Sewell (1989). Models of frontogenesis with the property $q = 0$ have been discussed by, for example, Hoskins and Bretherton (1972) and Purser and Cullen (1987).

The function $P(x_i, x_\alpha)$ in Fig. 19 is actually composed of straight lines, on each of which X_i is constant, and a string model of it can be constructed (Sewell, 1977, 1987). The convexified version in Fig. 20 of such $P(x_i, x_\alpha)$ is therefore only weakly convex. It may be more desirable physically, if q is to be positive, that $P(x_i, x_\alpha)$ be strictly jointly convex, as (97) would indicate. In meteorology, $q > 0$ expresses a stability property (Shutts and Cullen, 1987).

Chynoweth, Porter and Sewell (1988) used an example

$$R(X_i, X_\alpha) = \tfrac{1}{4} X_i^4 + X_i X_\alpha^2 + \tfrac{1}{40} X_i^2 + X_\alpha^2 \qquad (99)$$

(with $i = \alpha = 1$) of the parabolic umbilic to generate a piecewise smooth jointly convexified function $P(x_i, x_\alpha)$, having a crease representing a front across which the gradients of $P(x_i, x_\alpha)$ jump. Figs. 21 and 22 show corresponding parts of two such functions, from Chynoweth and Sewell (1989, Figs 7 and 8). The plan view of a discretized version of such $P(x_i, x_\alpha)$ is shown in Fig. 14. The implied q is not a constant.

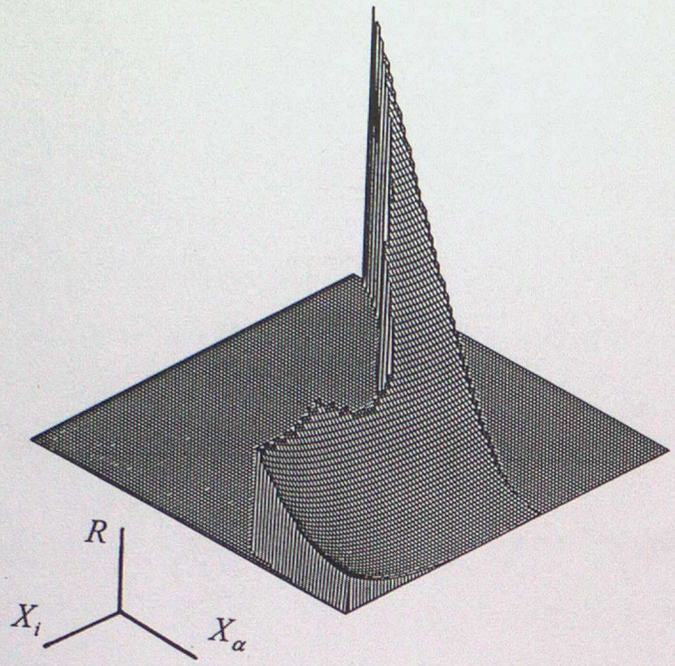


Fig. 21 Jointly convex parabolic umbilic $R(X_i, X_\alpha)$

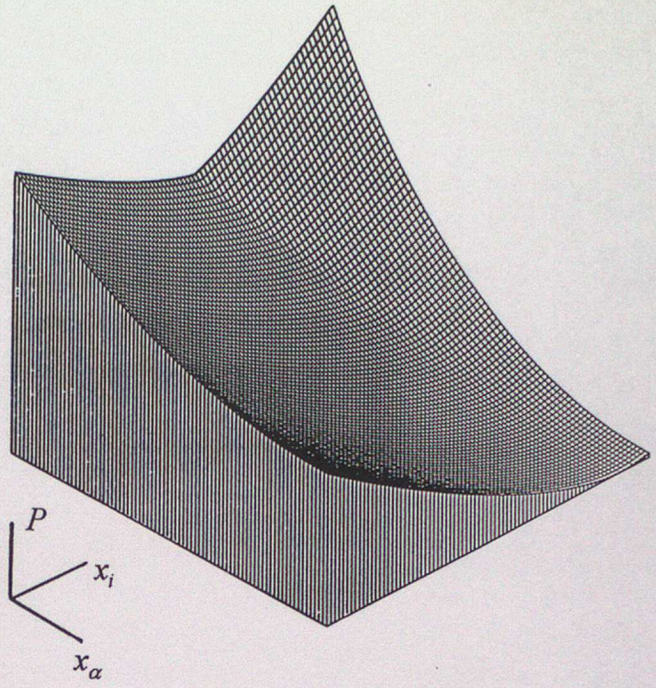


Fig. 22 Creased $P(x_i, x_\alpha)$ generated by parabolic umbilic

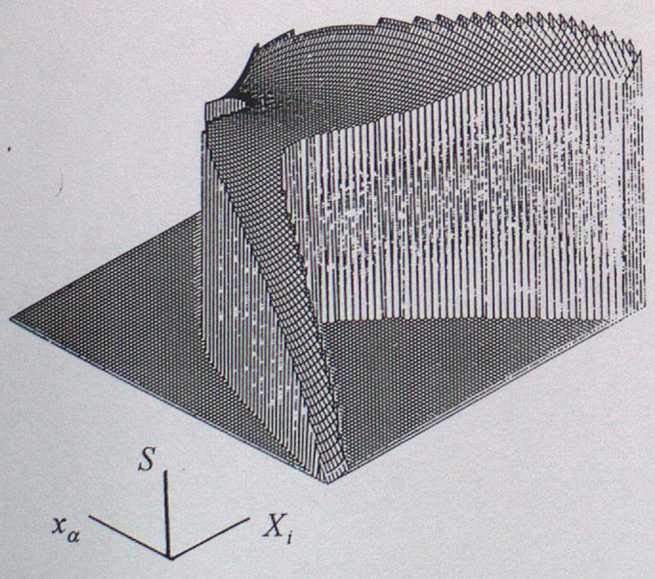


Fig. 23 Saddle function $S(X_i, x_\alpha)$ in equation (100)

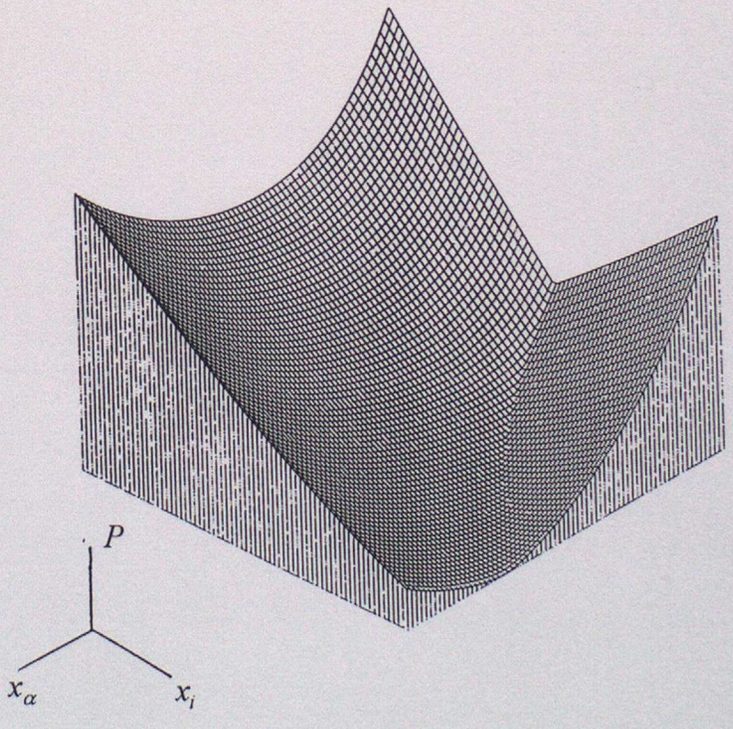


Fig. 24 Creased $P(x_i, x_\alpha)$ generated by saddle function with $q \equiv 1$.

Another starting point whose consequences were explored by Chynoweth and Sewell (1989, §§13 to 16) is a family of saddle functions $S(X_i, x_\alpha)$, again with $i = \alpha = 1$, but much wider than (98), and which include

$$S(X_i, x_\alpha) = \tfrac{1}{12} X_i^4 + \tfrac{1}{6} x_\alpha X_i^3 + \tfrac{1}{2}(x_\alpha - x_\alpha^2) X_i^2$$

$$+\left(\tfrac{5}{2} x_\alpha - \tfrac{1}{6} x_\alpha^3\right) X_i - \tfrac{1}{6} x_\alpha^3 + \tfrac{1}{12} x_\alpha^4 \ . \qquad (100)$$

This is an example of a case in which $q \equiv 1$, as may be verified directly from (97) because

$$\frac{\partial^2 S}{\partial X_i^2} = X_i^2 + x_\alpha X_i + x_\alpha - x_\alpha^2 = -\frac{\partial^2 S}{\partial x_\alpha^2} \ . \qquad (101)$$

In fact (100) was deduced from a much wider family of saddle functions associated with an assignable function $q(X_i, x_\alpha)$ and a family of explicit equations $x_i = x_i(x_\alpha)$ describing the plan view of the front ((54) of Chynoweth and Sewell, 1989). Fig. 23 shows the saddle function (100) over a domain which corresponds to the rectangular physical domain in x_i, x_α space, over which the jointly convex dual function $P(x_i, x_\alpha)$ exhibits the desired crease as shown in Fig. 24, whose plan view represents the front just mentioned. The corresponding

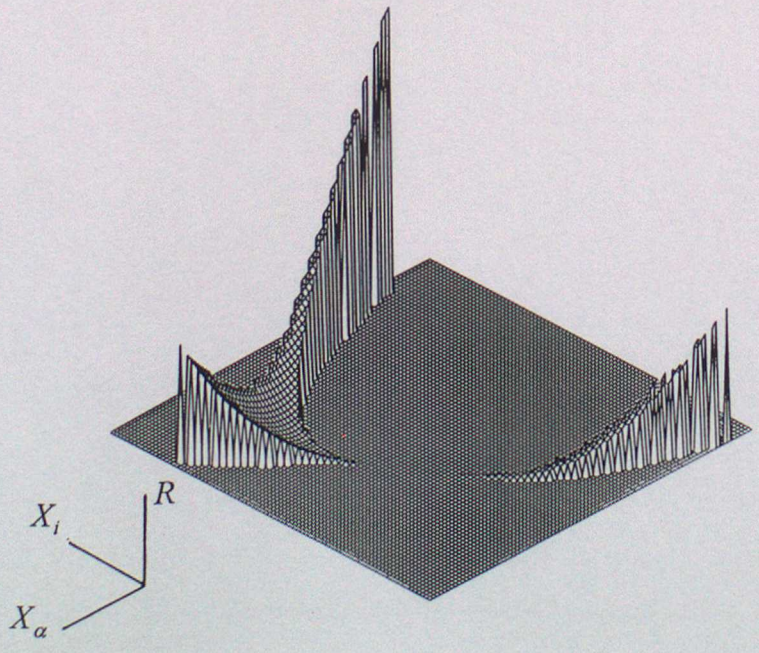


Fig. 25 Example of $R(X_i, X_\alpha)$ solving the Monge-Ampère equation (102).

jointly convex function $R(X_i, X_\alpha)$ illustrates a solution of the Monge-Ampère equation

$$\frac{\partial^2 R}{\partial(X_i, X_\alpha)} = 1 \ . \tag{102}$$

It is obtained by solving the cubic $-X_\alpha = \partial S/\partial x_\alpha$ needed to perform the Legendre transformation in (95). This $R(X_i, X_\alpha)$ is shown in Fig. 25. The jumps in the gradients $X_i = \partial P/\partial x_i$ and $X_\alpha = \partial P/\partial x_\alpha$ at the crease in Fig. 24, as given by $(94)_2$ and $(93)_1$ and based upon the starting point (100), are displayed in the functions $X_i(x_i, x_\alpha)$ and $X_\alpha(x_i, x_\alpha)$ for $i = \alpha = 1$ shown in Figs. 26 and 27. These jumps represent the discontinuities of velocity and temperature at the modelled front, as functions of the physical space coordinates in a vertical plane cross-section of the atmosphere.

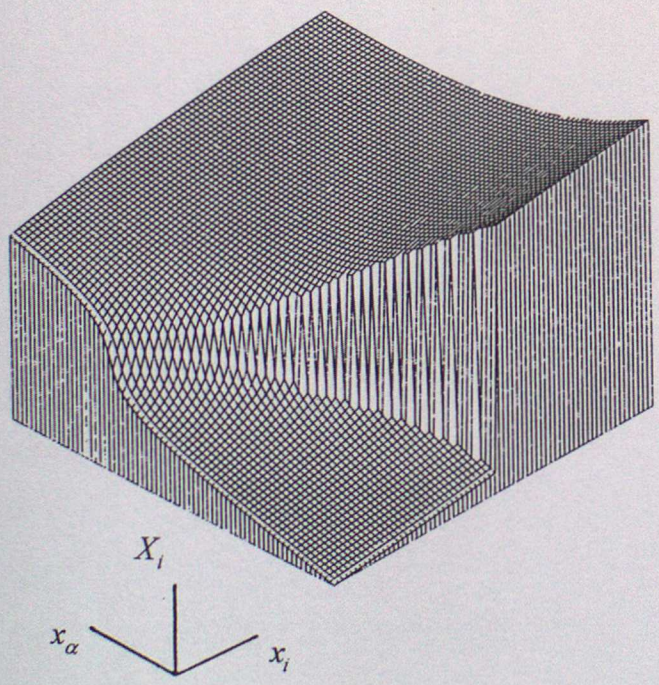


Fig. 26 Discontinuous $X_i(x_i, x_\alpha)$

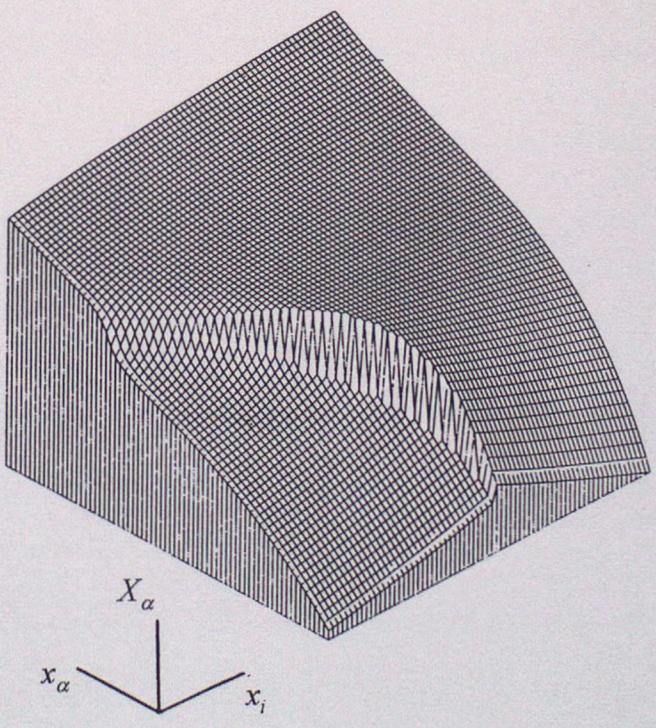


Fig. 27 Discontinuous $X_\alpha(x_i, x_\alpha)$

Other physical examples of quartets of Legendre transformations, for example in plasticity theory, thermodynamics and compressible fluid mechanics, are described by Sewell (1987, §§2.6, 2,7, 5.2, 5.3).

7. Flux of a jacobian

We record a result in the differential calculus, which is available more generally than may be widely appreciated. Let

$$b_i = b_i(x_j, t) \quad \text{with} \quad x_j = x_j(b_i, t) \tag{103}$$

be any invertible differentiable t-dependent mapping $\mathbb{R}^3 \leftrightarrow \mathbb{R}^3$ with jacobians

$$J = \frac{\partial(b_1, b_2, b_3)}{\partial(x_1, x_2, x_3)} \quad \text{and} \quad \frac{1}{J} = \frac{\partial(x_i)}{\partial(b_j)} . \tag{104}$$

Let

$$v_i = v_i(x_j, t)$$

be any other differentiable t-dependent mapping $\mathbb{R}^3 \to \mathbb{R}^3$, which may, but need not, represent a velocity. Let

$$D = \frac{\partial}{\partial t} + v_i \frac{\partial}{\partial x_i} \tag{105}$$

define an operator applicable to any function of the x_i and t.

Theorem 18

$$\frac{DJ}{J} = \frac{\partial D b_i}{\partial b_i} - \frac{\partial v_i}{\partial x_i} . \tag{106}$$

The proof given by Chynoweth and Sewell (1991, equation (A6)) admits no more than $3\, x_i$ (for other generic purposes we have used $n\, x_i$ in §2, and 2 or $1\, x_i$ in (84) and §6).

If x_i are cartesian eulerian particle position coordinates, t is time, and the v_i *are* particle velocity components, then $D = \overline{\dot{(\ )}}$, say, the time derivative following the particle. Then $v_i = \dot{x}_i$ and

$$\frac{\dot{J}}{J} = \frac{\partial \dot{b}_i}{\partial b_i} - \frac{\partial \dot{x}_i}{\partial x_i} \quad . \tag{107}$$

Several distinct examples of (107) are available. The b_i may or may not represent particle labels. When they do, they are also called lagrangian coordinates and $\dot{b}_i \equiv 0$. Mass balance then expresses J as the quotient of current to initial density in a general compressible medium, or of current to initial depth in incompressible shallow water theory. In both cases (107) has the structure of the continuity equation in a compressible medium.

In the following two examples the b_i are *not* particle labels. If $b_1 = x_1, b_2 = x_2$, but b_3 is the pseudo-height of Hoskins (1971), (107) becomes $\dot{J}/J = \partial \dot{b}_3/\partial x_3 - \partial \dot{x}_3/\partial x_3$ (Chynoweth and Sewell, 1991, equation (10)). Adiabatic flow of the atmosphere implies that *this* $\dot{J} = 0$, and the Boussinesq approximation then implies an incompressible version of the equation of continuity (see also Roulstone and Sewell, 1996, §4.2).

When the three b_i represent the geostrophic momenta X_i (for $i = 1,2$) together with X_α (for $\alpha = 1$) as given by $(93)_1$ and $(94)_2$, the J in (107) becomes the potential vorticity q in (97), and (107) can be rewritten as

$$\frac{\dot{q}}{q} = \frac{\partial \dot{X}_i}{\partial X_i} + \frac{\partial \dot{X}_\alpha}{\partial X_\alpha} - \left(\frac{\partial \dot{x}_i}{\partial x_i} + \frac{\partial \dot{x}_\alpha}{\partial x_\alpha} \right) \tag{108}$$

(Chynoweth and Sewell, 1991, equation (20); Roulstone and Sewell, 1996, equation (93)).

Whenever there is available a *particular* choice of b_i such that $\partial \dot{b}_i / \partial b_i = 0$, the corresponding J (call it b) satisfies $\dot{b}/b = -\partial \dot{x}_i / \partial x_i$ by (107), and the *general* b_i and J in (107) then satisfy

$$\dot{\left(\frac{J}{b}\right)} = \frac{J}{b}\frac{\partial \dot{b}_i}{\partial b_i} \quad . \tag{109}$$

Shallow water theory provides an example in which x_1 and x_2 are horizontal position coordinates, b is height, the general b_i are the geostrophic coordinates $X_i = \partial P/\partial x_i$, and the general J is the q of (97) without the x_α there.

If the subsequent dynamics enforces zero divergences, e.g. via theorems of Liouville type, it is (108) or (109) which establish conservation of q or of J/b following a particle.

8. Canonical transformations

The adjective 'canonical' is widely used in mechanics, but sometimes loosely. It may be prudent to ask what meaning is intended in each context. Sewell and Roulstone (1993) reviewed the concept of a canonical transformation, gave a variety of explicit examples, and noted the differences between several possible definitions of the transformation. Here we summarize those differences, in the simplest possible context, which offers a special clarity.

Again we use a neutral notation, not biased towards any particular context. We often use the same symbol to denote different differentiable functions which have the same value. Suppose that, for each value of a passive parameter t, a transformation from one pair of scalars x, y to another pair X, Y is expressed by

$$X = X(x,y,t) \quad , \quad Y = Y(x,y,t) \quad \text{with} \quad j = \frac{\partial(X,Y)}{\partial(x,y)} \quad . \tag{110}$$

The coordinates need not be cartesian, and the domains need not be the whole space. The transcriptions $x,y,X,Y \leftrightarrow p,q,P,Q$ can serve as a reminder of a classical mechanics interpretation in certain places.

Our *first definition* is that the transformation in (110) is canonical over any domain where the jacobian $j = 1$.

An example with t absent is $X = x^{1/2}\cos 2y$, $Y = x^{1/2}\sin 2y$ where $x \geq 0$. With cartesian coordinates this can represent the bending of an incompressible rectangular block in particle label (lagrangian) x, y space to part of a circular annulus in physical (eulerian) X, Y space (Green and Zerna, 1954, §3.11). Incompressibility in some form is a familiar consequence of a canonical transformation.

The condition $j \neq 0$ or $\pm\infty$ is sufficient for the inverse function theorem to guarantee a locally unique universe of $(110)_{1,2}$. Even when the functions in (110) are single valued, however, the global inverse of the transformation, which we write as

$$x = x(X,Y,t) \quad , \quad y = y(X,Y,t) \tag{111}$$

in terms of differentiable functions on the right, need not be single valued. An illustration is $x = X^2 + Y^2$, $y = \frac{1}{2}\arctan Y/X$ obtained from the example in the previous paragraph.

Where $j(\neq 0$ or $\pm\infty)$ is a constant for each t, the variables can be rescaled to achieve $j = 1$, but not where j depends on x or y or both.

Conformal transformations satisfy the Cauchy-Riemann equations. The question arises whether a transformation can be both canonical and conformal.

Theorem 19

The transformation in (110) is canonical (in the sense of $j = 1$) and conformal if and only if it has the particular orthogonal linear form

$$X = x\cos u(t) + y\sin u(t) \quad , \quad Y = -x\sin u(t) + y\cos u(t) \tag{112}$$

where $u(t)$ is any function of t alone.

Proof

The Cauchy-Riemann equations imply

$$j=\left(\frac{\partial X}{\partial x}\right)^2+\left(\frac{\partial X}{\partial y}\right)^2=\left(\frac{\partial Y}{\partial x}\right)^2+\left(\frac{\partial Y}{\partial y}\right)^2 .$$

Where this depends on t only we suppose the variables to have been rescaled to achieve $j=1$. The result is satisfied if and only if a function $u(x,y,t)$ exists such that

$$\frac{\partial X}{\partial x}=\cos u=\frac{\partial Y}{\partial y} \quad , \quad \frac{\partial X}{\partial y}=\sin u=-\frac{\partial Y}{\partial x} .$$

Commutivity of the second derivatives requires

$$\frac{\partial u}{\partial x}\cos u+\frac{\partial u}{\partial y}\sin u=0 \quad , \quad -\frac{\partial u}{\partial x}\sin u+\frac{\partial u}{\partial y}\cos u=0$$

and therefore $\partial u/\partial x=\partial u/\partial y=0$. Hence u can depend only on t, and (112) follows by integration.

□

Leaving aside conformal transformations now, additional hypotheses facilitate fresh and richer viewpoints, as the next results illustrate.

Theorem 20

In the neighbourhood of a point where $\partial Y/\partial x\neq 0$ or $\pm\infty$, the transformation (110) can be rewritten locally as

$$X=X(Y,y,t), \quad x=x(Y,y,t) \quad \text{with} \quad \frac{\partial X}{\partial y}=-j\frac{\partial x}{\partial Y} \tag{113}$$

in which $\partial x/\partial Y\neq 0$ or $\pm\infty$.

Proof

The hypothesis allows $(110)_2$ to be inverted as $x=x(Y,y,t)$, and inserted into $(110)_1$ and $(110)_2$ to give the new functions $X(Y,y,t)=X\big(x(Y,y,t),y,t\big)$ and

$Y = Y(x(Y,y,t), y, t)$ in terms, on the right, of the old ones of (110). The chain rule applied three times gives

$$\frac{\partial X}{\partial y} = \frac{\partial X}{\partial x}\frac{\partial x}{\partial y} + \frac{\partial X}{\partial y} \quad , \quad 1 = \frac{\partial Y}{\partial x}\frac{\partial x}{\partial Y} \quad , \quad 0 = \frac{\partial Y}{\partial x}\frac{\partial x}{\partial y} + \frac{\partial Y}{\partial y} \ ,$$

from which the result follows.

□

When the Maxwell relation

$$\frac{\partial X}{\partial y} + \frac{\partial x}{\partial Y} = 0 \tag{114}$$

associated with $(113)_1$ and $(113)_2$ is also available, it is equivalent to $j = 1$ under the extra hypotheses required by Theorem 20, because $(j-1)\,\partial x/\partial Y = 0$ is implied.

Legendre transformations expressible in terms of generating functions, having the structure of those displayed in Fig. 17, have a close link to canonical transformations. Fig. 28 shows such a quartet, and it possesses sufficient diversity to encapsulate the otherwise hidden 'anatomy' of a canonical transformation, as explained by the study of local hypotheses in Theorems 2 to 7 of Sewell and Roulstone (1993).

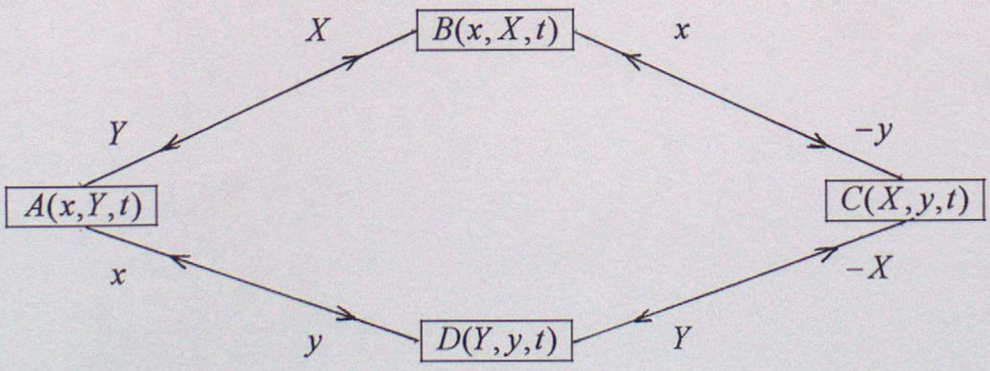


Fig. 28 Quartet of Legendre transformation

For example, a *second definition* could be that the transformation in (110) is canonical over any domain where at least one generating function exists, say $D(Y,y,t)$, with

$$X = -\frac{\partial D}{\partial Y} \quad , \quad x = \frac{\partial D}{\partial y} \quad \text{and} \quad \frac{\partial^2 D}{\partial Y \, \partial y} \neq 0 \, , \, \pm\infty \, . \tag{115}$$

The single hypothesis $(115)_3$ here is sufficient to allow (110) to be constructed locally, which establishes the applicability of the definition.

Theorem 21

The first derivatives of $X(x, y, t)$ and $Y(x, y, t)$ in (110) are related to the second derivatives of the generating function $D(Y, y, t)$ (denoted by suffixes here) by

$$\begin{aligned} \frac{\partial X}{\partial x} &= -\frac{D_{YY}}{D_{Yy}} \quad , \quad \frac{\partial X}{\partial y} = -D_{Yy} + \frac{D_{YY} \, D_{yy}}{D_{Yy}} \, , \\ \frac{\partial Y}{\partial x} &= \frac{1}{D_{Yy}} \quad , \quad \frac{\partial Y}{\partial y} = -\frac{D_{yy}}{D_{Yy}} \, . \end{aligned} \tag{116}$$

Proof

The hypothesis $D_{Yy} \neq 0$ or $\pm\infty$ in $(115)_3$ is sufficient to allow $x = x(Y, y, t)$ in $(115)_2$ to have inverse $Y = Y(x, y, t)$. Back substituting and differentiating with respect to x and y gives

$$1 = D_{yY} \frac{\partial Y}{\partial x} \quad \text{and} \quad 0 = D_{Yy} \frac{\partial Y}{\partial y} + D_{yy} \, ,$$

and hence $(116)_3$ and $(116)_4$. The same hypotheses allows $X = X(Y, y, t)$ in $(115)_1$ to be written $X = X\big(Y(x, y, t), y, t\big)$ whose partial derivatives are

$$\frac{\partial X}{\partial x} = -D_{YY} \frac{\partial Y}{\partial x} \quad \text{and} \quad \frac{\partial X}{\partial y} = -D_{YY} \frac{\partial Y}{\partial y} - D_{Yy}$$

from $(115)_1$, and hence $(116)_1$ and $(116)_2$.

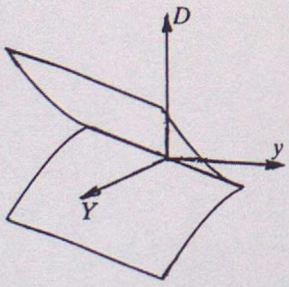
□

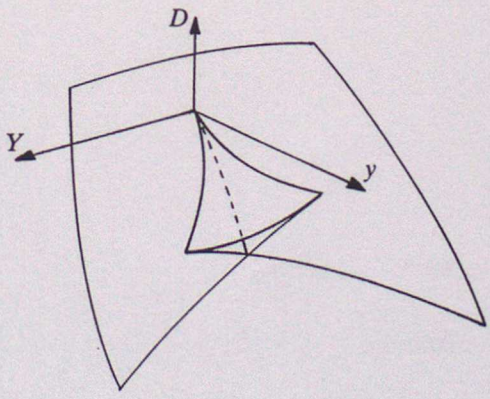
It follows directly from (116) that $j = 1$. Therefore our second definition of a canonical transformation implies our first one, but not vice versa. The two definitions are not

equivalent, and more hypotheses are inherent in the second one. The Maxwell relation (114) is necessary and sufficient for the local existence of $D(Y,y,t)$ satisfying $(115)_1$ and $(115)_2$.

Similar considerations relate to the other three generating functions displayed in Fig. 28. Sewell and Roulstone (1993) explain some of the singular features which can be associated with their existence.

The example $D(Y,y) = -\frac{1}{2}Y^2 \cot 2y$ generates the canonical transformation which describes the bending of a block, and the restriction $|y| < \frac{1}{4}\pi$ is needed if single-valuedness is required. The cusped example $D(Y,y) = \mp\frac{2}{3}(Y-y)^{3/2}$ shown in Fig. 29(a) generates the canonical transformation $X = x,\ Y = x^2 + y$. The conformal canonical transformation (112) is generated by, for example, $D(Y,y,t) = \frac{1}{2}(Y^2+y^2)\cot u - Yy\,\mathrm{cosec}\,u$ where $u \neq k\pi$ for any integer k.





(a) $D(Y,y) = \mp\frac{2}{3}(Y-y)^{3/2}$ (b) Swallowtail $D(Y,y)$

Fig. 29 Examples of multi-valued generating functions

A strong motivation for the study of canonical transformations is their role in hamiltonian mechanics, which can be illustrated by the next result. It is now convenient to think of the previously passive parameter t as time, and to denote total derivatives with respect to t by the superposed dot.

Theorem 22

Let $h(x,y,t)$ be a hamiltonian function associated with the hamiltonian pair of ordinary differential equations

$$\dot{x} = -\frac{\partial h}{\partial y} \quad , \quad \dot{y} = \frac{\partial h}{\partial x} \tag{117}$$

for determination of the functions $x(t)$ and $y(t)$. A canonical transformation in the sense of our second definition leaves the form of (117) invariant in the sense that it implies

$$\dot{X} = -\frac{\partial H}{\partial Y} \quad , \quad \dot{Y} = \frac{\partial H}{\partial X} \tag{118}$$

where $H(X,Y,t)$ is a function whose values are, for example,

$$H = h + \frac{\partial D}{\partial t} \ . \tag{119}$$

If t is absent from the canonical transformation, then $H = h$, i.e. the value of the hamiltonian is invariant as well.

Proof

Suppose that a new function with the property $\mathcal{H}(X,Y,t) = h(x,y,t)$ is associated with the transformation (110). Whether (110) is canonical or not, (117) and the chain rule imply

$$\dot{X} = -j\frac{\partial \mathcal{H}}{\partial Y} + \frac{\partial X}{\partial t} \quad , \quad \dot{Y} = j\frac{\partial \mathcal{H}}{\partial X} + \frac{\partial Y}{\partial t} \ .$$

A canonical transformation has $j = 1$ and, by the passive variable property in the Legendre transformation shown in Fig. 28 which satisfies $C + D = -XY$, we may write

$$X = -\frac{\partial D}{\partial Y} \quad , \quad Y = -\frac{\partial C}{\partial X} \quad \text{with} \quad \frac{\partial D}{\partial t} + \frac{\partial C}{\partial t} = 0 \ .$$

Then

$$\dot{X} = -\frac{\partial \mathcal{H}}{\partial Y} - \frac{\partial}{\partial Y}\left(\frac{\partial D}{\partial t}\right) \quad , \quad \dot{Y} = \frac{\partial \mathcal{H}}{\partial X} + \frac{\partial}{\partial X}\left(\frac{\partial D}{\partial t}\right) .$$

□

Hamiltonian methods are a strong current feature in geophysical fluid dynamics. Numerical integration algorithms which preserve hamiltonian structure are called symplectic integrators.

If a generating function such as $D(Y,y,t)$ with the properties (115) exists, then its first differential at each fixed t is

$$dD = -XdY + xdy \ , \tag{120}$$

unambiguously on any single-valued branch of $D(Y,y,t)$. It follows that

$$\oint XdY = \oint xdy \ , \tag{121}$$

where the circuit integral is evaluated around any closed curve in the Y, y domain for the given t. The circuit is, by implication, lifted onto the single-valued branch under consideration.

A possible *third definition* of a canonical transformation can be expressed in terms of the invariance of a circuit integral as follows. The transformation (110) is canonical if at least one of the invariance properties (121) or

$$\oint XdY = -\oint ydx \ , \quad \oint YdX = \oint ydx \ , \quad \oint YdX = -\oint xdy \tag{122}$$

holds. The last three are associated with the other generating functions $A(x,Y)$, $B(x,X)$, $C(X,y)$ respectively in Fig. 28. According to Arnold (1989, §§44E and 45A), a definition of this type is the generally accepted one, and implies others which he quotes (cf. Schutz, 1980). However, as explained by Sewell and Roulstone (1993), such a definition is not equivalent to the other ones, and it is subject to qualifications associated with the fact that one or more of the generating functions can easily possess multivaluedness or domain degeneracy, which would need to be assessed in advance before a circuit integral can be associated with a specific

branch. For example, smooth and continuous $D(Y,y)$ can be double-valued as in Fig. 29(a), or swallowtailed as in Fig. 29(b). The latter example has the general form of an inverted version of $P(x_i, x_\alpha)$ in Fig. 19(b), and comparison with (98) shows that it could result from choosing $C(X,y) = -X^4 - y\,X^2$ in Fig. 28 and applying the Legendre transformation. Thermodynamics supplies other examples of swallowtailed functions associated with canonical transformations, such as the Gibbs free enthalpy function for the van der Waals fluid (see Fig. 12(b)), as explained by Sewell and Roulstone (1993, §5(c)) and Sewell (1987, §5.2).

The Poisson bracket of two differentiable functions $F(X,Y,t)$ and $G(X,Y,t)$ is defined to be

$$[F,G] = \frac{\partial F}{\partial X}\frac{\partial G}{\partial Y} - \frac{\partial F}{\partial Y}\frac{\partial G}{\partial X} \tag{123}$$

for each value of t. A particular example is that $[F,F]$ is the jacobian of $F(X,Y,t)$. Suppose the transformation (110), whether canonical or not according to some definition, induces two new functions $f(x,y,t) = F(X,Y,t)$ and $g(x,y,t) = G(X,Y,t)$. Results of the following type (cf. Theorem 8 of Sewell and Roulstone, 1993) establish the connection between invariance of the Poisson bracket and definitions of the canonical transformation.

Theorem 23

Under the hypothesis of Theorem 20

$$[F,G]-[f,g] = \left(\frac{\partial f}{\partial x}\frac{\partial G}{\partial X} - \frac{\partial F}{\partial X}\frac{\partial g}{\partial x}\right)\frac{\partial x}{\partial Y}(j-1)\ . \tag{124}$$

Proof

We use $(113)_2$ to define new functions $\bar{f}(Y,y,t) = f\big(x(Y,y,t),y,t\big)$ and $\bar{g}(Y,y,t) = g\big(x(Y,y,t),y,t\big)$, whence

$$\frac{\partial \bar{f}}{\partial Y} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial Y} = \frac{\partial F}{\partial X}\frac{\partial X}{\partial Y} + \frac{\partial F}{\partial Y} \quad , \quad \frac{\partial \bar{g}}{\partial y} = \frac{\partial g}{\partial x}\frac{\partial x}{\partial y} + \frac{\partial g}{\partial y} = \frac{\partial G}{\partial X}\frac{\partial X}{\partial y}$$

by the chain rule. Hence

$$\left(\frac{\partial F}{\partial Y} + \frac{\partial F}{\partial X}\frac{\partial X}{\partial Y}\right)\frac{\partial G}{\partial X} = \frac{\partial f}{\partial x}\frac{\partial G}{\partial X}\frac{\partial x}{\partial Y} \quad , \quad \frac{\partial f}{\partial x}\left(\frac{\partial g}{\partial y} + \frac{\partial g}{\partial x}\frac{\partial x}{\partial y}\right) = \frac{\partial f}{\partial x}\frac{\partial G}{\partial X}\frac{\partial X}{\partial y} .$$

Similarly, by interchanging the roles of f and g, and of F and G,

$$\left(\frac{\partial G}{\partial Y} + \frac{\partial G}{\partial X}\frac{\partial X}{\partial Y}\right)\frac{\partial F}{\partial X} = \frac{\partial g}{\partial x}\frac{\partial F}{\partial X}\frac{\partial x}{\partial Y} \quad , \quad \frac{\partial g}{\partial x}\left(\frac{\partial f}{\partial y} + \frac{\partial f}{\partial x}\frac{\partial x}{\partial y}\right) = \frac{\partial g}{\partial x}\frac{\partial F}{\partial X}\frac{\partial X}{\partial y} .$$

Subtracting this pair from the previous pair gives

$$[F,G] = \left(\frac{\partial F}{\partial X}\frac{\partial g}{\partial x} - \frac{\partial f}{\partial x}\frac{\partial G}{\partial X}\right)\frac{\partial x}{\partial Y} \quad , \quad [f,g] = \left(\frac{\partial f}{\partial x}\frac{\partial G}{\partial X} - \frac{\partial F}{\partial X}\frac{\partial g}{\partial x}\right)\frac{\partial X}{\partial y} .$$

Hence

$$[f,g] - [F,G] = \left(\frac{\partial f}{\partial x}\frac{\partial G}{\partial X} - \frac{\partial F}{\partial X}\frac{\partial g}{\partial x}\right)\left(\frac{\partial X}{\partial y} + \frac{\partial x}{\partial Y}\right) .$$

Using $(113)_3$ delivers (124).

□

It is evident that when (124) holds, the extra hypothesis $j = 1$ implies

$$[F,G] = [f,g] , \tag{125}$$

but that (125) does not imply $j = 1$ because the first of the three factors on the right of (124) might be zero. This fact, and similar features associated with the other three ways of pairing the variables according to whether (110) is assumed to have $\partial Y/\partial y$ or $\partial X/\partial y$ or $\partial X/\partial x \neq 0$ or $\pm\infty$, indicates that (125) is not equivalent to $j = 1$.

A possible *fourth definition* that (110) be a canonical transformation is that (125) holds for *every* pair of functions $F(X,Y,t)$ and $G(X,Y,t)$. This is preferred by some authors such as Sudarshan and Mukunda (1974) and Woodhouse (1987). Formulae such as (124),

valid where $\partial Y/\partial x \neq 0$ or $\pm\infty$, inform a comparison of this definition with the previous ones.

Our *fifth definition* is that a transformation is canonical if it leaves the form of hamilton's equations invariant. Theorem 22 shows how this is implied by our second definition, but does not assert the converse, i.e. it does not say that invariance of the form of hamilton's equations implies the existence of generating functions.

An important particular case in which such a converse can be established is conveyed by the following result. Here, invariance of hamilton's equations is shown to imply $j = 1$, which is necessary for the existence of generating functions by Theorem 21. It follows from the next Theorem that at least some generating functions exist, with a property representing (119).

Theorem 24

Suppose that a transformation having the properties

$$y = y(Y,t) \text{ with } \frac{\partial y}{\partial Y} \neq 0 \text{ or } \pm\infty, \text{ and } \dot{y} = \frac{\partial y}{\partial Y}\dot{Y} + \frac{\partial y}{\partial t}, \tag{126}$$

the second equation being a consequence of the first, converts the function on the left of

$$l(\dot{y},y,t) = L(\dot{Y},Y,t) \tag{127}$$

into the function on the right, without changing its value, as indicated. Then the two equations

$$\frac{d}{dt}\left(\frac{\partial l}{\partial \dot{y}}\right) - \frac{\partial l}{\partial y} = 0 \quad \text{and} \quad \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{Y}}\right) - \frac{\partial L}{\partial Y} = 0 \tag{128}$$

are equivalent, i.e. when one holds, so does the other. It follows that introduction of the new variables

$$x = \frac{\partial l}{\partial \dot{y}} \quad \text{and} \quad X = \frac{\partial L}{\partial \dot{Y}} \tag{129}$$

allows equations (128) to be rewritten in equivalent hamiltonian forms (117) and (118) in terms of hamiltonian functions

$$h(x,y,t) = x\dot{y} - l \quad \text{and} \quad H(X,Y,t) = X\dot{Y} - L \,, \tag{130}$$

if we assume that the standard Legendre transformations required are non-singular. The transformation relating x, y, X, Y, t has the unsymmetrical implicit form

$$y = y(Y,t) \quad , \quad \frac{\partial h}{\partial x} = \frac{\partial y}{\partial Y}\frac{\partial H}{\partial X} + \frac{\partial y}{\partial t} \tag{131}$$

and the property $j = 1$. It also has the property

$$H - h = X\frac{\partial H}{\partial X} - x\frac{\partial h}{\partial x} \,. \tag{132}$$

It follows that at least $A(x,Y,t)$ and $C(X,y,t)$ exist, such that

$$H - h = -\frac{\partial A}{\partial t} = -\frac{\partial C}{\partial t} \,. \tag{133}$$

Proof

The chain rule applied to (127) with (126) leads to

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{Y}}\right) - \frac{\partial L}{\partial Y} = \left(\frac{d}{dt}\left(\frac{\partial l}{\partial \dot{y}}\right) - \frac{\partial l}{\partial y}\right)\frac{\partial y}{\partial Y} + \left(\frac{d}{dt}\left(\frac{\partial y}{\partial Y}\right) - \frac{\partial \dot{y}}{\partial Y}\right)\frac{\partial l}{\partial \dot{y}}$$

and

$$\frac{d}{dt}\left(\frac{\partial y}{\partial Y}\right) = \frac{\partial \dot{y}}{\partial Y} = \frac{\partial^2 y}{\partial Y^2}\dot{Y} + \frac{\partial^2 y}{\partial t\,\partial Y} \,.$$

This establishes the equivalence in (128), and hence that of the hamiltonian versions (117) and (118). The form of hamilton's equations is thereby left invariant by a transformation (110) which has the particular form expressed implicitly by (131), and alternatively by

$$\frac{\partial H}{\partial X} = \left(\frac{\partial h}{\partial x} - \frac{\partial y}{\partial t}\right)\frac{\partial Y}{\partial y} \quad \text{with} \quad Y = Y(y,t) \,.$$

Differentiating this with respect to x gives

$$\frac{\partial^2 H}{\partial X^2}\frac{\partial X}{\partial x} = \frac{\partial^2 h}{\partial x^2}\frac{\partial Y}{\partial y} \ ,$$

so the jacobian in $(110)_3$ is

$$j = \frac{\partial X}{\partial x}\frac{\partial Y}{\partial y} = \frac{\partial^2 h/\partial x^2}{\partial^2 H/\partial X^2}\left(\frac{\partial Y}{\partial y}\right)^2 \ .$$

Second derivatives with respect to the active variables in a Legendre transformation (see, e.g. Sewell, 1987, Chapter 2) have the properties

$$\frac{\partial^2 l}{\partial \dot{y}^2}\frac{\partial^2 h}{\partial x^2} = 1 \quad \text{and} \quad \frac{\partial^2 L}{\partial \dot{Y}^2}\frac{\partial^2 H}{\partial X^2} = 1 \ .$$

Also, from (127)

$$\frac{\partial L}{\partial \dot{Y}} = \frac{\partial l}{\partial \dot{y}}\frac{\partial \dot{y}}{\partial \dot{Y}} = \frac{\partial l}{\partial \dot{y}}\frac{\partial y}{\partial Y} \quad \text{so} \quad \frac{\partial^2 L}{\partial \dot{Y}^2} = \frac{\partial^2 l}{\partial \dot{y}^2}\left(\frac{\partial y}{\partial Y}\right)^2 \ .$$

Combining these results gives $j = 1$, and (132) follows from (127), (130), $(117)_2$ and $(118)_2$. Our assumptions above that $\partial y/\partial Y$, $\partial^2 L/\partial \dot{y}^2$ and $\partial^2 L/\partial \dot{Y}^2$ are finite and non-zero imply the same for $\partial Y/\partial y$, $\partial^2 h/\partial x^2$, $\partial^2 H/\partial X^2$ and $\partial X/\partial x$. From this follow the Maxwell relations

$$\frac{\partial X}{\partial x} = \frac{\partial y}{\partial Y} \quad \text{and} \quad \frac{\partial x}{\partial X} = \frac{\partial Y}{\partial y}$$

in place of (114), and hence the existence of generating potentials $A(x,Y,t)$ and $C(X,y,t)$ such that

$$X = \frac{\partial A}{\partial Y} \ , \ y = \frac{\partial A}{\partial x} \quad \text{and} \quad x = -\frac{\partial C}{\partial y} \ , \ Y = -\frac{\partial C}{\partial X}$$

(cf. Theorems 2 and 3 of Sewell and Roulstone, 1993). Then (133) is an alternative to (119), using the property of derivatives with respect to passive variables in a Legendre transformation.

□

An example of Theorem 24 is

$$y = u(t)Y \quad \text{with} \quad l(\dot{y}, y) = \tfrac{1}{2} m\dot{y}^2 - v(y) \tag{134}$$

with any constant $m \neq 0$ and any differentiable functions $u(t) \neq 0$ and $v(y)$. This has

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{Y}}\right) - \frac{\partial L}{\partial Y} = u\left[\frac{d}{dt}\left(\frac{\partial l}{\partial \dot{y}}\right) - \frac{\partial l}{\partial y}\right], \quad j = 1, \quad H - h = -\dot{u}\,x\,Y = -\frac{\dot{u}}{u^2} Xy \tag{135}$$

and the associated canonical transformation (110) is, from (131),

$$X = u\,x, \quad Y = \frac{y}{u}, \quad \text{with} \quad A(x, Y, t) = u\,x\,Y, \quad C(X, y, t) = -\frac{Xy}{u}. \tag{136}$$

9. Central orbits and semi-geostrophic theory

The partial differential equations of semi-geostrophic rotating fluid theory have provided stimulus for the deployment of several of the ideas explained in this article. That theory stems from the work of Hoskins (1975), Fjortoft (1962) and Eliassen (1948), beginning with constant Coriolis parameter f, and the Hamiltonian-based approximations of Salmon (1983, 1985, 1988) for variable f. Quite detailed physical and mathematical arguments are required to justify the approximations involved in replacing Newton's Second Law by semi-geostrophic theory, as is clear, for example, from the articles by Hoskins (1975) and Cullen (1997) in this Volume. Further mathematical properties of such models are the object of current research, for example by Roulstone and Sewell (1996, 1997), McIntyre and Roulstone (1996) and Theiss (1997).

An elementary prototype of the semi-geostrophic approximation is explored in this Section, in terms of ordinary differential equations, which allows some very explicit conclusions.

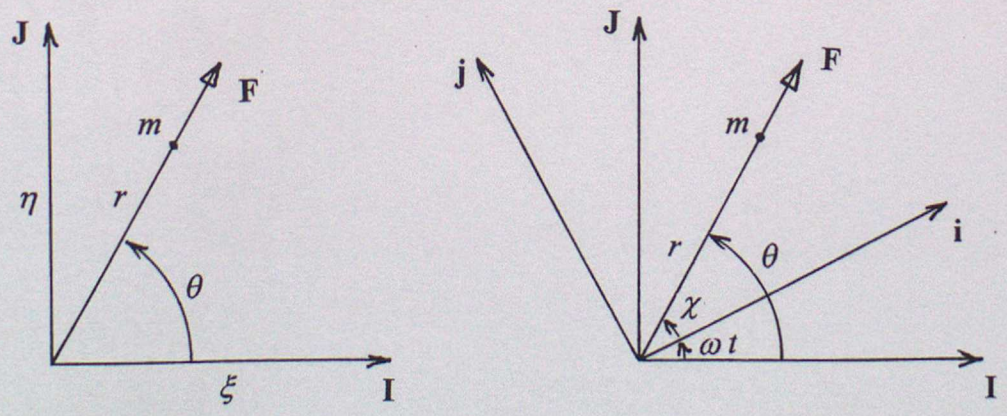
Newtonian dynamics of the plane motion of a single particle under a central force, viewed from an inertial frame, consists of a pair of second order ordinary differential

equations, whose solution will involve four constants of integration. Under the inverse square law of force, for example, the solutions are conic sections. The same problem can also be viewed from a reference frame which is uniformly rotating relative to the inertial one. In this frame the semi-geostrophic approximation can be introduced, which replaces Newton's equations by a pair of first order ordinary differential equations, whose solution will involve only two constants of integration. In itself that is a radical simplification from the previous four constants.

We shall prove that, in general, the semi-geostrophic solutions are all circles, under the same inverse square law of central force which provides the much richer family of Newtonian solutions, namely all the rotating conics (as viewed from the rotating frame), whether they be ellipses, hyperbolae or parabolae.

Thus this problem allows us to watch explicitly how the semi-geostrophic approximation imposes a rather drastic filtering of solutions.

The inertial frame can be regarded as fixed in space in the plane of the page, and spanned by cartesian coordinates ξ, η with unit vectors $\mathbf{I}$, $\mathbf{J}$, or by polar coordinates r, θ. A particle of mass m has position vector $\mathbf{r} = \xi\mathbf{I} + \eta\mathbf{J}$, and is acted upon by a central force



(a) Inertial frame (b) Rotating frame

Fig. 30 Particle under central force

$\mathbf{F} = F\mathbf{r}/r$ as shown in Fig. 30(a). Here F is a given function $F(r)$ only, so that we may write $F = -dV/dr$ where $V(r)$ is another function interpretable as potential energy. Newton's Second Law states that $m\ddot{\mathbf{r}} = F\mathbf{r}/r$ or

$$\ddot{r} - r\dot{\theta}^2 = -\frac{1}{m}\frac{dV}{dr} \quad , \quad \frac{d}{dt}(r^2\dot{\theta}) = 0 \ . \tag{137}$$

Two of the four constants of integration can be written

$$\tfrac{1}{2}m\left(\dot{r}^2 + \frac{M^2}{r^2}\right) + V = h \quad , \quad r^2\dot{\theta} = M \tag{138}$$

where h is the total energy and $m\,M$ is the moment of momentum about the origin. We deviate here from standard notation, which uses h instead of M.

Introduce another reference frame, with the same fixed origin as the inertial one, but rotating with constant anticlockwise spin ω relative to it, and spanned by cartesian coordinates x, y with unit vectors $\mathbf{i}, \mathbf{j}$ as shown in Fig. 30(b) after time t, so that $\mathbf{r} = x\mathbf{i} + y\mathbf{j}$ also. The consequent transformation equations

$$\xi = x\cos\omega t - y\sin\omega t \quad , \quad \eta = x\sin\omega t + y\cos\omega t \tag{139}$$

and their inverses and time derivatives facilitate the introduction of a modified potential function

$$\phi(r) = \frac{V(r)}{m} - \tfrac{1}{2}\omega^2 r^2 \tag{140}$$

which allow Newton's Second Law to be rewritten, with $r = \left(x^2 + y^2\right)^{1/2}$, as

$$\ddot{x} - 2\omega\dot{y} + \frac{\partial\phi}{\partial x} = 0 \quad , \quad \ddot{y} + 2\omega\dot{x} + \frac{\partial\phi}{\partial y} = 0 \tag{141}$$

in terms of the rotating frame, for any central force law. In passing we notice that (139) is an example of (112) in which $u(t) = -\omega t$.

In the particular case of the inverse square law,

$$F = -\frac{m\mu}{r^2} \quad , \quad V = -\frac{m\mu}{r} \quad , \quad \phi = -\frac{\mu}{r} - \tfrac{1}{2}\omega^2 r^2 \tag{142}$$

with constant μ, and $\mu > 0$ implies attraction. The solutions of (137) with (142) are well known to be fixed conics viewed from the inertial frame, so the corresponding solutions of (141) will be rotating conics viewed from the rotating frame.

It is informative to notice how the four constants of integration arise in these inverse square law Newtonian solutions. First we recall a standard result (cf. Pars, 1953, §14.4).

Theorem 25

The general solution of $m\ddot{\mathbf{r}} = F\mathbf{r}/r$ with the inverse square law (142) lies on the curve

$$r(1 - A\cos\theta - B\sin\theta) = C \tag{143}$$

where $C = M^2/\mu$, and A and B are constants of integration. In terms of the inertial cartesian coordinates (143) implies

$$(1-A^2)\xi^2 - 2AB\xi\eta + (1-B^2)\eta^2 - 2AC\xi - 2BC\eta - C^2 = 0 \ , \tag{144}$$

which confirms that (143) is a conic section.

Proof

Newton's Second Law can be rewritten, using $(138)_2$ and $(142)_1$, as

$$\ddot{\xi} = -\frac{\mu}{M}\dot{\theta}\cos\theta \quad , \quad \ddot{\eta} = -\frac{\mu}{M}\dot{\theta}\sin\theta \ . \tag{145}$$

Integrating once gives

$$\dot{\xi} = \frac{\mu}{M}(B - \sin\theta) \quad , \quad \dot{\eta} = \frac{\mu}{M}(\cos\theta - A) \tag{146}$$

where A and B are arbitrary constants. We obtain (143) by inserting (146) into $M = r(\dot{\eta}\cos\theta - \dot{\xi}\sin\theta)$. Then (144) is obtained by squaring (143).

□

Only three of the four constants h, M, A, B of the motion appear in (143), and (143) does not yet express the orbit solutions as space coordinate functions of time.

Theorem 26

Suppose the initial conditions are that, at $t = 0$,

$$r = a > 0 \quad , \quad \theta = \alpha \quad , \quad \dot{\xi} = u \quad , \quad \dot{\eta} = v \tag{147}$$

where a, α, u, v are given constants. The previous four constants of the motion can be expressed in terms of them as

$$h = \tfrac{1}{2}m(u^2 + v^2) - m\frac{\mu}{a} \quad , \quad M = a(v\cos\alpha - u\sin\alpha) \tag{148}$$

$$A = \cos\alpha - \frac{vM}{\mu} \quad , \quad B = \sin\alpha + \frac{uM}{\mu} \, . \tag{149}$$

The proof is immediate from (138) and (146). Thus any given choice of the initial position and velocity a, α, u, v uniquely specifies the constants h, M, A, B.

A parabolic orbit will serve to give an illustration of an orbital solution with simple algebra.

Theorem 27

The apsidal initial conditions $\alpha = u = 0$ with $v = (2\mu/a)^{1/2}$ in (147) induce a parabolic orbit

$$\xi = a(1 - \tau^2) \quad , \quad \eta = 2a\tau \tag{150}$$

where the parameter τ is related to the real time t by

$$\tau(t) = p - \frac{1}{p} \quad \text{where} \quad p(t) = \left[\frac{3t}{2k} + \left(1 + \frac{9t^2}{4k^2}\right)^{1/2}\right]^{1/3} \tag{151}$$

and $k = a^{3/2}(2/\mu)^{1/2}$.

Proof

The stated conditions imply $h = B = 0,\ A = -1,\ M = (2\mu a)^{1/2},\ C = 2a,$ so that $\eta^2 + 4a\xi = 4a^2$ from (144). This can be written parametrically as (150) which, when inserted into either of (146), gives $k(1+\tau^2)\dot{\tau} = 1$. Integrating with the choice that $\tau = 0$ when $t = 0$, to be consistent with the initial conditions, gives $t = k(\tau + \frac{1}{3}\tau^3)$. This cubic is monotonic, has no extrema or stationary points, and its solution is (151).

□

The literature (e.g. Pars, 1953, §14.9) shows that elliptical or hyperbolic orbits, respectively, result from apsidal initial conditions $\alpha = u = 0$ with $v^2 < 2\mu/a$ or $v^2 > 2\mu/a$.

The complete parabolic orbital solution of Newton's Second Law with the stated initial conditions is obtained by inserting (151) into (150). The same solution as viewed from the rotating frame, i.e. of (140), is then obtained from (139) as the rotating parabola

$$x(t) = \xi(t)\cos\omega t + \eta(t)\sin\omega t\,, \quad y(t) = -\xi(t)\sin\omega t + \eta(t)\cos\omega t\,. \tag{152}$$

The semi-geostrophic approximation to Newton's Second Law in this context is obtained by starting with (141), defining the geostrophic velocity u_g, v_g and Coriolis parameter f by

$$u_g = -\frac{1}{f}\frac{\partial\phi}{\partial y}\quad,\quad v_g = \frac{1}{f}\frac{\partial\phi}{\partial x}\quad,\quad f = 2\omega\,, \tag{153}$$

and substituting $\dot{u}_g, \dot{v}_g$ in place of $\ddot{x}, \ddot{y}$ in (141), giving

$$-\frac{1}{f}\frac{\dot{\partial\phi}}{\partial y} + \frac{\partial\phi}{\partial x} - f\dot{y} = 0\quad,\quad \frac{1}{f}\frac{\dot{\partial\phi}}{\partial x} + \frac{\partial\phi}{\partial y} + f\dot{x} = 0\,. \tag{154}$$

This direct procedure replaces the pair of second order equations (141), whose solution will normally require *four* assignable constants as we have seen, with a pair of first order equations (154).

We may expect that only *two* assignable constants will be required to solve the latter. The question arises how this difference between the number of required constants affects a comparison of the two systems of equations, and in particular how a solution of the second system may be regarded as an approximation to the solution of the first system.

With constant $f = 2\omega$, as we have assumed in this Section, we can use (140) to introduce the auxiliary function

$$\lambda(r) = \frac{1}{f} + \frac{1}{f^3 r}\frac{d\phi}{dr} = \frac{3}{4f} - \frac{F}{mrf^3} \quad \text{so that} \quad \frac{d\phi}{dr} = f^2 r(f\lambda - 1) \ . \tag{155}$$

The semi-geostrophic equations (154) then become

$$-\frac{d}{dt}(\lambda y) = (1 - f\lambda)x \ , \quad \frac{d}{dt}(\lambda x) = (1 - f\lambda)y \tag{156}$$

and therefore

$$(\dot{y} - xf)\lambda + y\frac{d\lambda}{dr}\dot{r} + x = 0 \ , \quad (\dot{x} + yf)\lambda + x\frac{d\lambda}{dr}\dot{r} - y = 0 \ . \tag{157}$$

Since $r^2 = x^2 + y^2$ and $r\dot{r} = x\dot{x} + y\dot{y}$, this is a pair of first order linear ordinary differential equations but with variable coefficients.

The next result expresses the drastic filtering which is enforced by the semi-geostrophic approximation in this context.

Theorem 28

The solutions of the semi-geostrophic equations are, in general, circular orbits as seen from the rotating frame provided the force law is such that $\lambda(r) \neq c/r$ for any constant c. If

b is the radius of the circle, the orbital solution in terms of the polar angle χ in the rotating frame (see Fig. 30(b)) is

$$r = b \quad , \quad \chi = \left(f - \frac{1}{\lambda(b)} \right) t \ . \tag{158}$$

This reduces to a single equilibrium point at $r = b, \chi = 0$ in the rotating frame if $\lambda(b) = 1/f$, i.e. if $d\phi/dr = 0$ at $r = b$. This point will appear as a circular orbit in the inertial frame.

If $\lambda(r) = c/r$ the force law will be the linear one

$$F = \tfrac{3}{4} m f^2 r - m f^3 c \ . \tag{159}$$

The problem is then underdetermined, and arbitrarily many orbits are possible.

The inverse square law of force admits the circular orbits only.

Proof

Multiplying $(157)_1$ by y and $(157)_2$ by x and adding implies

$$r\dot{r} \frac{d}{dr}(\lambda r) = 0 \ . \tag{160}$$

Where $r \neq 0$ this implies $\dot{r} = 0$ or $\lambda r = c$.

Thus if $\lambda r \neq c$, the only orbits are circles in general. If the initial conditions associated with (154), (156) or (157) are that $x = b$, $y = 0$ at $t = 0$, the circular solution path $x = b\cos\chi$, $y = b\sin\chi$ inserted into (157) implies the orbital solution (158). Equilibrium in the rotating frame occurs when the particular values $\lambda(b) = 1/f$ and therefore $F = -mb\omega^2$ apply, which in the case of the inverse square law would require $\mu = b^3\omega^2$.

When $\lambda r = c$, writing $x = r\cos\chi$, $y = r\sin\chi$ converts (156) into

$$(c\dot{\chi} + r - fc)\cos\chi = 0 \quad , \quad (c\dot{\chi} + r - fc)\sin\chi = 0 \ , \tag{161}$$

which have arbitrarily many solutions. They are motions $r = c(f - \dot{\chi})$ with *any* chosen function $\chi(t)$, or equilibrium points $r = fc$ with any fixed χ.

Since (159) is not the inverse square law, the last part of the Theorem is confirmed.

□

The semi-geostrophic orbits are therefore, in general, always circles, whereas the Newtonian orbits include (under the inverse square law) the much richer set of rotating conic sections, including ellipses, hyperbolae and parabolae. There is, however, a very simple relation between the sets of orbits in the two cases, which can be illustrated as follows.

Theorem 29

The envelope of all the inverse square law Newtonian orbits which start from an apse with $\alpha = u = 0$ at $r = a$ with fixed $a > 0$ and arbitrary given v is the semi-geostrophic circular orbit of radius a.

Proof

With arbitrary v, as distinct from the particular v chosen in Theorem 27, and the stated apsidal conditions, the conic section solutions (143) may be written

$$G(A) = 0 \quad \text{where} \quad G(A) = r - A\xi + a(A-1) , \tag{162}$$

because $M = av$, $A = 1 - av^2/\mu$, $B = 0$, $C = a(1-A)$. The envelope of these conics is obtained by eliminating the parameter A from (162) and $dG/dA = 0$, which is $\xi = a$, thus giving $r = a$. In Theorem 28 we can choose $b = a$ here.

□

The language used by dynamical systems theorists, for example in partial differential equations, is that the geostrophic approximation to Newtonian dynamics express a geostrophic balance, between Coriolis and pressure gradient forces, which picks out a 'slow manifold' of solutions, within the larger Newtonian phase space where unwanted 'fast' solutions also lie. In our simple example it seems that the circular envelope represents the slow manifold. It needs to be recalled, however, that in the semi-geostrophic theory of meteorology, the potential vorticity is often regarded as assigned, which leads to a Monge-Ampère equation (97), with the hamiltonian (91) (see also (190)) regarded as unknown *a priori*, in contrast to the known (140) with (138) in this Section.

10. Hamiltonian structures

Although complete solutions to the ordinary differential equations for the Newtonian problem (under the inverse square law) and its semi-geostrophic replacement are obtained, for a single particle, in Theorems 25 and 28, it is worth studying their hamiltonian structure briefly in this Section. This is because they act as a prototype for the corresponding partial differential equations in shallow water theory and atmospheric dynamics, whose complete solutions are not known.

First we record the generalization of Theorem 24 to any number of dimensions, in the standard notation of analytical mechanics. This result is quite general, and not restricted to central orbit theory.

Theorem 30

Suppose that a transformation having the properties

$$q_i = q_i(Q_j,t) \quad \text{with} \quad \frac{\partial(q_i)}{\partial(Q_j)} \neq 0 \text{ or } \pm\infty, \quad \text{and} \quad \dot{q}_i = \frac{\partial q_i}{\partial Q_j}\dot{Q}_j + \frac{\partial q_i}{\partial t}, \tag{163}$$

with $i,j = 1, \ldots, n,$ the second system of equations being a consequence of the first, converts the function on the left of

$$l(\dot{q}_i,q_i,t) \;=\; L(\dot{Q}_i,Q_i,t) \tag{164}$$

into the function on the right, without changing its value, as indicated. Then the form of lagrange's equations is invariant in the sense that

$$\frac{d}{dt}\left(\frac{\partial l}{\partial \dot{q}_i}\right) - \frac{\partial l}{\partial q_i} \;=\; 0 \quad \text{and} \quad \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{Q}_i}\right) - \frac{\partial L}{\partial Q_i} \;=\; 0 \tag{165}$$

are equivalent, i.e. when one holds, so does the other. The form of the corresponding hamilton's equations is also invariant in the sense that the new variables

$$p_i = \frac{\partial l}{\partial \dot{q}_i} \quad \text{and} \quad P_i = \frac{\partial L}{\partial \dot{Q}_i} \tag{166}$$

convert (165) into

$$\dot{p}_i = -\frac{\partial h}{\partial q_i}, \quad \dot{q}_i = \frac{\partial h}{\partial p_i} \quad \text{and} \quad \dot{P}_i = -\frac{\partial H}{\partial Q_i}, \quad \dot{Q}_i = \frac{\partial H}{\partial P_i} \tag{167}$$

where

$$h(p_i, q_i, t) = p_i \dot{q}_i - l \quad \text{and} \quad H(P_i, Q_i, t) = P_i \dot{Q}_i - L \ . \tag{168}$$

The transformation can be rewritten

$$\frac{\partial H}{\partial P_i} = \frac{\partial Q_i}{\partial q_j}\left(\frac{\partial h}{\partial p_j} - \frac{\partial q_j}{\partial t}\right) \quad , \quad Q_i = Q_i(q_j, t) \tag{169}$$

as a particular but implicit example of

$$P_i = P_i(p_j, q_j, t) \quad , \quad Q_i = Q_i(p_j, q_j, t) \ . \tag{170}$$

It has the property

$$H - h = P_i \frac{\partial H}{\partial P_i} - p_i \frac{\partial h}{\partial p_i} \ . \tag{171}$$

Also, there exist generating functions

$$A(p_i, Q_i, t) = p_i \, q_i(Q_j, t) \quad \text{and} \quad C(P_i, q_i, t) = -P_i \, Q_i(q_j, t) \tag{172}$$

such that

$$P_i = \frac{\partial A}{\partial Q_i}, \quad q_i = \frac{\partial A}{\partial p_i} \quad \text{and} \quad p_i = -\frac{\partial C}{\partial q_i}, \quad Q_i = -\frac{\partial C}{\partial P_i} \tag{173}$$

and

$$H - h = -\frac{\partial A}{\partial t} = -\frac{\partial C}{\partial t} \ . \tag{174}$$

Proof

The results down to (171) might be regarded as standard, and the proof follows the first steps in that of Theorem 24, using the jacobian property in (163), and the Legendre transformations.

The proofs of $(173)_2$ and $(173)_4$ are immediate. The other gradients of (172) are

$$\frac{\partial A}{\partial Q_i} = p_j \frac{\partial q_j}{\partial Q_i} \quad \text{and} \quad \frac{\partial C}{\partial q_i} = -P_j \frac{\partial Q_j}{\partial q_i} .$$

From $(163)_2$, (164) and (166) we have

$$P_i = \frac{\partial L}{\partial \dot{Q}_i} = \frac{\partial l}{\partial \dot{q}_j} \frac{\partial \dot{q}_j}{\partial \dot{Q}_i} = p_j \frac{\partial q_j}{\partial Q_i}, \quad \text{and} \quad p_i = P_j \frac{\partial Q_j}{\partial q_i}$$

similarly, which proves $(173)_1$ and $(173)_3$.

Then $(169)_1$ and its counterpart, and (171), give

$$H - h = -p_j \frac{\partial q_j}{\partial t} = P_j \frac{\partial Q_j}{\partial t} ,$$

which proves (174).

□

Theorem 30 allows us to say that the transformation (169), regarded as a particular and implicit example of (170), is a *canonical transformation* in the sense of both our second and fifth definitions in §8. It illustrates how those definitions, at least, readily generalize to higher dimensions. Neither the q_i nor the Q_i need be cartesian coordinates, but either can be such.

A specific illustration is available in §9. The transformation (139) is an example of (163) with the transcriptions $\xi, \eta, x, y \leftrightarrow q_i, q_2, Q_1, Q_2$, and it is conformal with $\partial(q_i)/\partial(Q_j) = 1$, being an example of Theorem 19 with $u(t) = -\omega t$, as we remarked after (141). It participates in a canonical transformation whose other half, corresponding to $(163)_2$, is

$$\begin{aligned} \dot{\xi} &= (\dot{x} - \omega y)\cos\omega t - (\dot{y} + \omega x)\sin\omega t , \\ \dot{\eta} &= (\dot{x} - \omega y)\sin\omega t + (\dot{y} + \omega x)\cos\omega t . \end{aligned} \tag{175}$$

The classical lagrangian, as seen from the inertial frame, is

$$
\begin{aligned}
l(\dot{\xi},\dot{\eta},\xi,\eta) &= \tfrac{1}{2}m(\dot{\xi}^2+\dot{\eta}^2)-V \\
&= \tfrac{1}{2}m(\dot{x}^2+\dot{y}^2)-m\phi+m\omega(x\dot{y}-y\dot{x}) = L(\dot{x},\dot{y},x,y)\,,
\end{aligned}
\tag{176}
$$

expressed on the right as seen from the rotating frame, using (140) and $r^2=\xi^2+\eta^2=x^2+y^2$. The momenta as viewed from the inertial and rotating frames are $p_1=m\dot{\xi}$, $p_2=m\dot{\eta}$, $P_1=m(\dot{x}-\omega y)$, $P_2=m(\dot{y}+\omega x)$ respectively. The two hamiltonians are

$$
\begin{aligned}
&h(p_1,p_2,\xi,\eta)=\tfrac{1}{2m}(p_1^2+p_2^2)+V\,, \\
&H(P_1,P_2,x,y)=\tfrac{1}{2m}(P_1+m\omega y)^2+\tfrac{1}{2m}(P_2-m\omega x)^2+m\phi\,.
\end{aligned}
\tag{177}
$$

The equivalence and form invariance of governing equations under the transformation (139) with (175) can now conveniently be expressed as follows, for the motion of a single particle under a central force.

Theorem 31

Newton's Second Law (137), viewed from the inertial frame in Fig. 30(a), is expressible equivalently as

$$
m\ddot{\xi} = -\frac{\xi}{r}\frac{dV}{dr} \quad , \quad m\ddot{\eta} = -\frac{\eta}{r}\frac{dV}{dr} \tag{178}
$$

and also as the lagrangian form $(165)_1$ with $(176)_1$, and again as the hamiltonian form $(167)_1$ and $(167)_2$ with $(177)_1$.

The same problem, viewed from the rotating frame in Fig. 30(b), is expressible equivalently as (141), and also as the lagrangian form $(165)_2$ with $(176)_2$, and again as the hamiltonian form $(167)_3$ and $(167)_4$ with $(177)_2$.

The generating functions (172) in this example are

$$
\begin{aligned}
A(p_1,p_2,x,y,t) &= (p_1x+p_2y)\cos\omega t+(p_2x-p_1y)\sin\omega t\,, \\
C(P_1,P_2,\xi,\eta,t) &= -(P_1\xi+P_2\eta)\cos\omega t+(P_2\xi-P_1\eta)\sin\omega t\,.
\end{aligned}
\tag{179}
$$

Proof

From the standpoint now reached, the proof is by direct verification using the formulae already established.

□

For example, the idea that $(179)_1$ *generates*, by differentiation in $(173)_1$, the part $(169)_1$ of the canonical transformation, is illustrated by the result

$$P_1 = p_1 \cos\omega t + p_2 \sin\omega t\,, \qquad P_2 = -p_1 \sin\omega t + p_2 \cos\omega t \tag{180}$$

of that procedure. It can be seen that (180) is equivalent to (175). Also, (180) is conformal, like (139). Thus *both* parts of (169) are conformal, separately, in this example.

The formula (174) is illustrated in this example via

$$H - h = \omega(P_1 y - P_2 x) = \omega(p_2 \xi - p_1 \eta) = -\frac{\partial A}{\partial t} = -\frac{\partial C}{\partial t}\,. \tag{181}$$

We observe that t does not appear explicitly either in the two lagrangians (176) or in the two hamiltonians (177) in the present example, but it *does* appear in the canonical transformation (139) with (180) which converts one system into the other.

We saw in §9 that the semi-geostrophic approximation replaces a pair of second order differential equations (141) by a pair of first order equations (154). Alternatively, the approximation can be viewed as replacing one hamiltonian system of four first order equations, described in two alternative but equivalent forms in Theorem 31, by another hamiltonian system of two first order equations. Here we obtain the latter by using the ideas in §§5 and 6.

Theorem 32

The semi-geostrophic equations (154) are equivalent to the hamiltonian equations

$$\dot{X} + f\frac{\partial \Psi}{\partial Y} = 0 \quad , \quad \dot{Y} - f\frac{\partial \Psi}{\partial X} = 0 \tag{182}$$

where X and Y are defined by

$$X = x + \frac{1}{f^2}\frac{\partial \phi}{\partial x} \quad , \quad Y = y + \frac{1}{f^2}\frac{\partial \phi}{\partial y} \tag{183}$$

and the hamiltonian function $f\,\Psi(X,Y)$ is actually a function of $R = (X^2 + Y^2)^{1/2}$ only, as defined in (190) below. Furthermore

$$\dot{X} = u_g \quad , \quad \dot{Y} = v_g \ . \tag{184}$$

Proof

Equations (183) are an example of $(84)_1$ with (85), which can be written as (86) in the form

$$X = \frac{\partial P}{\partial x} \ , \ Y = \frac{\partial P}{\partial y} \quad \text{where} \quad P(x,y) = \tfrac{1}{2}r^2 + \tfrac{1}{f^2}\phi(r) \tag{185}$$

in terms of (140). To approach the inversion of the Legendre transformation, rewrite (185) as

$$X = x\,f\lambda(r) \quad , \quad Y = y\,f\lambda(r) \tag{186}$$

in terms of the auxiliary function introduced in (155). Then

$$R = r\,f\,\lambda(r) \ \text{ with solution } \ r = \Lambda(R) \ , \tag{187}$$

say, so that the inverse of (186) is

$$x = \frac{X\Lambda(R)}{R} = \frac{\partial S}{\partial X} \quad , \quad y = \frac{Y\Lambda(R)}{R} = \frac{\partial S}{\partial Y} \tag{188}$$

following $(87)_1$. Here the Legendre dual function is

$$S(X,Y) \ = \ xX + yY - P \ = \ R\Lambda - \tfrac{1}{2}\Lambda^2 - \tfrac{1}{f^2}\phi(\Lambda) \ , \tag{189}$$

which is a function of R alone. In terms of it, following (88), we can construct the function

$$\Psi(X,Y) \ = \ \tfrac{1}{2}R^2 - S \ = \ \tfrac{1}{2}\big(R - \Lambda(R)\big)^2 + \tfrac{1}{f^2}\phi\big(\Lambda(R)\big) \tag{190}$$

which depends on R alone. Differentiation of (185) and (190) shows that

$$\dot{X}+f\frac{\partial\Psi}{\partial Y}=\dot{x}+\frac{1}{f^2}\dot{\overline{\frac{\partial\phi}{\partial x}}}+\frac{1}{f}\frac{\partial\phi}{\partial y}\quad,\quad \dot{Y}-f\frac{\partial\Psi}{\partial X}=\dot{y}+\frac{1}{f^2}\dot{\overline{\frac{\partial\phi}{\partial y}}}-\frac{1}{f}\frac{\partial\phi}{\partial x}\quad, \tag{191}$$

which proves the Theorem. It also follows from (141) and (185) that (184) holds.

□

In the special case of the inverse square law (142), (187) is

$$R=\frac{3r}{4}+\frac{\mu}{f^2r^2}\quad. \tag{192}$$

We conclude by noting a generalization of Theorem 31. Working in the variables associated with the rotating frame of Fig. 30(b), define the function

$$L(\dot{x},\dot{y},x,y)=\tfrac{1}{2}m(\dot{x}^2+\dot{y}^2)-m\phi+mu\dot{x}+mv\dot{y}\quad, \tag{193}$$

where $u(x, y)$ and $v(x, y)$ are any pair of functions of x and y which satisfy

$$\frac{\partial v}{\partial x}-\frac{\partial u}{\partial y}=2\omega\quad. \tag{194}$$

Thus u and v can be thought of as the components of any vector field which has the associated vorticity 2ω.

For given ω, there will be many solutions of (194) because, for example, $u(x, y)$ can be chosen arbitrarily, and then $v(x, y)$ can be obtained by integration, which introduces an arbitrary function of y. A simple example of a family of functions satisfying (194) is

$$u=(c-\omega)y+a\quad,\quad v=(c+\omega)x+b \tag{195}$$

with arbitrary constants a, b, c. If $a=b=c=0$, (193) becomes $(176)_2$, but not otherwise.

Theorem 33 below will justify the interpretation of (193) as an *augmented* lagrangian. If we define the augmented momenta

$$\pi_1=\frac{\partial L}{\partial\dot{x}}=m(\dot{x}+u)\quad,\quad \pi_2=\frac{\partial L}{\partial\dot{y}}=m(\dot{y}+v) \tag{196}$$

we can use the Legendre transformation to construct an augmented hamiltonian function

$$H(\pi_1,\pi_2,x,y) = \frac{(\pi_1 - mu)^2}{2m} + \frac{(\pi_2 - mv)^2}{2m} + m\phi \tag{197}$$

having values $H = \pi_1\dot{x} + \pi_2\dot{y} - L$.

Theorem 33

The augmented lagrangian equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) - \frac{\partial L}{\partial x} = 0 \quad , \quad \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{y}}\right) - \frac{\partial L}{\partial y} = 0 \tag{198}$$

are equivalent to the augmented hamiltonian equations

$$\dot{\pi}_1 = -\frac{\partial H}{\partial x} \quad , \quad \dot{\pi}_2 = -\frac{\partial H}{\partial y} \quad , \quad \dot{x} = \frac{\partial H}{\partial \pi_1} \quad , \quad \dot{y} = \frac{\partial H}{\partial \pi_2} \quad , \tag{199}$$

and both (198) and (199) are equivalent to the Newtonian equations (141) viewed from the rotating frame, and therefore also to (178) viewed from the inertial frame.

Proof

It is enough to note, by direct verification from (193) and (194), that

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) - \frac{\partial L}{\partial x} = m\left(\ddot{x} - 2\omega\dot{y} + \frac{\partial \phi}{\partial x}\right) \quad , \quad \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{y}}\right) - \frac{\partial L}{\partial x} = m\left(\ddot{y} + 2\omega\dot{x} + \frac{\partial \phi}{\partial y}\right) . \tag{200}$$

□

This result offers an equivalent version of the Newtonian equations in terms of an augmented lagrangian (193) which does *not*, in general, have the same value as the original lagrangian (176). We might say that the latter hypothesis, whose general form is (164), has been replaced by the constraint (194). Variational principles associated with Theorem 33, but in shallow water theory, have been given by Salmon (1983) and Roulstone and Sewell (1996).

When considering changing the value of a lagrangian, it may be verified that, in the generality of Theorem 30, the total time derivative $\dot{n}$ of *any* function $n(q_i,t)$ is a function $\dot{n}(\dot{q}_i,q_i,t)$, linear in the $\dot{q}_i$, which satisfies lagrange's equations $(165)_1$ (cf. Theorem 24). In other words, if l satisfies lagrange's equations, so does $l+\dot{n}$. In the context of (193) with (194), any function $n(x,y)$ can generate an addition $\dot{n}=\dot{x}\,\partial n/\partial x+\dot{y}\,\partial n/\partial y$ to (193) whose implied additions to u and v have zero extra vorticity, and therefore still leave (194) satisfied.

References

Arnold, V.I. 1989 *Mathematical Methods of Classical Mechanics*, Second edition. Springer-Verlag, New York.

Carathéodory, C. 1982 *Calculus of Variations and Partial Differential Equations of the First Order*. Second (revised) English edition. Chelsea, New York.

Chynoweth, S., Porter D. and Sewell, M.J. 1988 The parabolic umbilic and atmospheric fronts. *Proc. R. Soc. Lond.* A **419**, 337-362.

Chynoweth, S. and Sewell, M.J. 1989 Dual variables in semi-geostrophic theory. *Proc. R. Soc. Lond.* A **424**, 155-186.

Chynoweth, S. and Sewell, M.J. 1990 Mesh duality and Legendre duality. *Proc. R. Soc. Lond.* A **428**, 351-377.

Chynoweth, S. and Sewell, M.J. 1991 A concise derivation of the semi-geostrophic equations, *Q. Jl. R. Met. Soc.* **117**, 1109-1128.

Coles, P. 1990 Understanding recent observations of the large-scale structure of the universe. Nature **346**, 446.

Collins, I.F. and Houlsby, G.T. 1997 Applications of thermomechanical principles to the modelling of geotechnical materials. *Proc. R. Soc. Lond.* A **453**, 1975-2001.

Cullen, M.J.P. 1997 New mathematical developments in atmosphere and ocean dynamics, and their application to computer simulations. Meteorological Office Forecasting Research Division Scientific Paper No. 48.

Cullen, M.J.P. and Purser, R.J. 1984 An extended Lagrangian theory of semi-geostrophic frontogenesis. *J. Atmos. Sci.* **41**, 1477-1497.

Delannay, B. 1934 Sur la sphere vide. Izvestia Akademii Nauk SSSR, Mathematical and Natural Sciences Division, no. 6, pp. 793-800.

Eliassen, A. 1948 The quasi-static equations of motion with pressure as independent variable. *Geofys. Publ.* **17**, 1-44.

Fjortoft, R. 1962 On the integration of a system of geostrophically balanced prognostic equations. *Proc. Intern. Symp. Numerical Weather Prediction*, Meteor. Soc. Japan, Tokyo, pp. 153-159.

Green A.E. and Zerna, W. 1954 Theoretical Elasticity. Clarendon Press. Oxford.

Hill, R. 1987 Constitutive dual potentials in classical plasticity. *J. Mech. Phys. Solids* **35**, 23-33.

Hoskins, B.J. 1971 Atmospheric frontogenesis models: some solutions. *Q. Jl. R. Met. Soc.* **97**, 139-153.

Hoskins, B.J. 1975 The geostrophic momentum approximation and the semi-geostrophic equations. *J. Atmos. Sci.* **32**, 233-242.

Hoskins, B.J. and Bretherton, F.P. 1972 Atmospheric frontogenesis models: mathematical formulation and solution. *J. Atmos. Sci.* **29**, 11-37.

McIntyre, M.E. and Roulstone, I. 1996 Hamiltonian balanced models: constraints, slow manifolds and velocity splitting. Meteorological Office Forecasting Research Division Scientific Paper No. 41.

Noble, B. 1964 Complementary variational principles for boundary value problems I: basic principles, with an application to ordinary differential equations. Report 473, Mathematics Research Center, University of Wisconsin.

Pars, L. 1953 Introduction to Dynamics. Cambridge University Press.

Poston, T. and Stewart, I.N. 1978 Catastrophe Theory and its Applications, Pitman, London.

Purser, R.J. 1993 Contact transformations and Hamiltonian dynamics in generalized semi-geostrophic theories. *J. Atmos. Sci.* **50**, 1449-1468.

Purser, R.J. 1996 Legendre-transformable semi-geostrophic theories. Isaac Newton Institute Preprint NI96014.

Purser, R.J. and Cullen, M.J.P. 1987 A duality principle in semigeostrophic theory. *J. Atmos. Sci.* **44**, 3449-3468.

Roulstone, I. and Sewell, M.J. 1996 Potential vorticities in semi-geostrophic theory. *Q.Jl. R. Met. Soc.* **122**, 983-992.

Roulstone, I. and Sewell, M.J. 1997 The mathematical structure of theories of semi-geostrophic type. *Phil. Trans. R. Soc. Lond.* A **355**, 1-29.

Rubtsov, V.N. and Roulstone, I. 1997 Examples of quaternionic and Kähler structures in Hamiltonian models of nearly geostrophic flow. *J. Phys. A: Math. Gen.* **30**, L63-L68.

Salmon, R. 1983 Practical use of Hamilton's principle. *J. Fluid Mech.*, **132**, 431-444.

Salmon, R. 1985 New equations for nearly geostrophic flow. *J. Fluid Mech.* **153**, 461-477.

Salmon, R. 1988 Semi-geostrophic theory as a Dirac bracket projection. *J. Fluid. Mech.* **196**, 345-358.

Schutz, B. 1980 *Geometrical Methods of Mathematical Physics*. Cambridge University Press.

Sewell, M.J. 1963 On reciprocal variational principles for perfect fluids. *J. Math. Mech.*, **12**, 495-504.

Sewell, M.J. 1966 On the connexion between stability and the shape of the equilibrium surface. *J. Mech. Phys. Solids*, **14**, 203-230.

Sewell, M.J. 1977 On Legendre transformations and elementary catastrophes. *Math. Proc. Camb. Phil. Soc.*, **82**, 147-163.

Sewell, M.J. 1978a A pure catastrophe machine. *Bull. Inst. Maths. Applics.* **14**, 212-214.

Sewell, M.J. 1978b On Legendre transformations and umbilic catastrophes. *Math. Proc. Camb. Phil. Soc.* **83**, 273-288.

Sewell, M.J. 1982 Legendre transformations and extremum principles. Pp. 563-605 of *Mechanics of Solids, The Rodney Hill 60th Anniversary Volume*, edited by H.G. Hopkins and M.J. Sewell, Pergamon Press, Oxford.

Sewell, M.J. 1987 *Maximum and Minimum Principles. A unified approach, with applications*. Cambridge University Press.

Sewell, M.J. 1990 Demonstrations of buckling by bifurcation and snapping. Pp. 315-345 of *Elasticity: Mathematical Methods and Applications. The Ian Sneddon 70th Birthday Volume*. Edited by G. Eason and R.W. Ogden. Ellis Horwood Ltd., Chichester.

Sewell, M.J. 1997 *Mathematics Masterclasses - Stretching the Imagination*. Oxford University Press.

Sewell, M.J. and Porter, D. 1980 Constitutive surfaces in fluid mechanics. *Math. Proc. Camb. Phil. Soc.* **88**, 517-546.

Sewell, M.J. and Roulstone, I. 1993 Anatomy of the canonical transformation. *Phil. Trans. R. Soc. Lond.* **345**, 577-598.

Sewell, M.J. and Roulstone, I. 1994 Families of lift and contact transformations. *Proc. R. Soc. Lond.* A **447**, 493-512.

Shutts, G.J., Booth, M. and Norbury, J. 1988 A geometric model of balanced, axisymmetric flows with embedded penetrative convection. *J. Atmos. Sci.* **45**, 2609-2621.

Shutts, G.J. and Cullen, M.J.P. 1987 Parcel stability and its relation to semi-geostrophic theory. *J. Atmos. Sci.* **44**, 1318-1330.

Sudarshan, E.C.G. and Mukunda, N. 1974 *Classical Dynamics. A modern perspective*. Wiley, New York.

Theiss, J. 1997 Hamiltonian balanced models. Department of Applied Mathematics and Theoretical Physics Report, Cambridge University.

Voronoi, A. 1908 Nouvelles applications des parametres continue à la théorie des formes quadratiques. *J. Reine Angew. Math.* **134**, 198-287.

Woodhouse, N.M.J. 1987 *Introduction to Analytical Dynamics*. Oxford University Press.

Zeeman, E.C. 1977 *Catastrophe Theory. Selected papers* 1972-77. Addison-Wesley, Reading, Mass.